ESSAYS ON THE ECONOMIC TRANSITION TO ADULTHOOD: MIGRATION, FERTILITY AND EDUCATION DECISIONS AMONG YOUNG WOMEN IN SUB-SAHARA AFRICA

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Catalina Herrera Almanza

August 2014

BIOGRAPHICAL SKETCH

Catalina Herrera Almanza holds a B.A. in Economics from the Universidad de Los Andes in Bogota, Colombia, and a M.A. in Economics from the same university. Prior to her doctoral studies at Cornell University, Catalina worked for Colombian government and academic institutions as well as multilateral organizations. She was an advisor at the Public Policy Evaluation Directorate in Colombia’s Ministry of National Planning, a research assistant at Fedesarrollo, the largest Colombian think-tank, Universidad de los Andes and the Colombian Central Bank.

Later on, Catalina worked as a Junior Professional Associate at the World Bank in Washington DC, conducting research on the migration and poverty linkages in Nicaragua and Morocco. She also worked as a consultant for the Office of Evaluation and Oversight at the Inter-American Development Bank (IADB) in education projects for the Caribbean region. During her doctoral studies at Cornell, she worked in rural Guatemala as a field coordinator of the impact evaluation of Heifer, an animal donation program.

Through this professional research experience as well as her field work experiences in Colombia, Guatemala and in Madagascar, as part of her doctoral thesis, Catalina has consolidated her research interests in development economics, with a primary focus on the microeconometric analysis of population, health and education issues in Sub Sahara Africa and Latin America.

At the time of writing, Catalina has accepted a joint appointment as an Assistant Professor in the Departments of Economics and in the International Affairs Program at Northeastern University, Boston (M.A).

To my parents, Jorge and Lucero, and my sister Ana Maria

To my fiancé, Shady

ACKNOWLEDGMENTS

I am most grateful to the members of my PhD Special Committee for their outstanding guidance during this dissertation process. First of all, I would like to thank David E. Sahn, the chair of my PhD dissertation committee, for his wise advising and mentoring through the dissertation and doctoral training. His discussions were highly stimulating and fostered critical thinking in the implementation of research projects. I am grateful to George Jakubson, member of my special committee, for his Econometrics training in an intuitive rather than mechanical fashion. I am also grateful to Ravi Kanbur, member of my special committee, for his advice on the public policy insights of my research projects.

I am thankful to Harivelo Rajemison and his team at the Madagascar's Institute of National Statistics with whom I worked to complete the data cleaning process of the household surveys used in these research projects.

This dissertation was partially supported by the Hewlett Foundation/IIE Dissertation Fellowship in Population, Reproductive Health, and Economic Development, the Cornell's Enaudi research travel grants and the Cornell Population Center grants.

I am deeply thankful to my fiancé, Shady Atallah, for his intellectual contribution and encouragement to this dissertation as well as his unconditional support. We have walked this PhD road side by side since we met in 2009 until now when we are graduating together. I am grateful to my friends, Sharmini Radakrishnan, Maria Fernanda Rosales and Camilo Garcia for their feedback and discussions. Finally, I thank my parents and sister for their continuous support and encouragement.

TABLE OF CONTENTS

CHAPTER 1

DETERMINANTS OF INTERNAL MIGRATION AMONG SENEGALESE YOUTH

Abstract

We analyze the socio-economic determinants of youth decision to internally migrate in Senegal. Young people undertake mostly rural-to-rural and urban-to-urban migrations and over half of them are temporary migrants. Using multinomial logit models, we estimate the role of household and community characteristics during childhood in later youth migration decisions. We find that these determinants are heterogeneous by gender and destination. The higher the fathers' education the more (less) likely are their daughters to move to urban (rural) areas. Young individuals, who spend their childhood in better off households, are more likely to move to urban areas. Also, the presence of younger siblings increases the propensity of moving to rural areas. Access to primary schools during childhood decreases the likelihood of migrating to urban areas for both men and women.

1.1 Introduction

Internal migration, mostly composed of young adults and the poor, constitutes the largest flow of people in developing countries (UNDP 2009). Recent empirical evidence has focused on analyzing the determinants as well as the impacts of international migration. The study of internal flows has been far more limited than international migration, partly due to the lack of reliable data as well as the fact that internal migration is far less of a political issue. The research on internal mobility patterns, including the main drivers and outcomes, is nonetheless relevant because it will provide information to policymakers to enhance the benefits of internal migration and manage its risks and costs.

The recent literature has pointed out the importance of family and social factors in the decisions of young adults to migrate (World Bank 2007). In developing countries, where households face labor and financial market constraints, migration can be a strategy to diversify income sources and cope with risks, compensating in some cases for the absence of insurance markets (Rosenzweig and Stark 1989; Stark 1991; Potick and Kuhn 2006; Giles 2007). Families might encourage younger members to migrate, both sons and daughters, not only because they have higher earnings potenti-al, but also because they are more likely to remit money (Taylor 2001). Furthermore, family and socioeconomic circumstances during childhood can influence his/her probability of migrating later in life (Abramitzky et al. 2012).

This paper addresses the following question: What are the socio-economic determinants of the decisions to internally migrate in Senegal among young adults? And are these factors differentiated by gender? In particular, we analyze whether the decision to migrate is influenced by individual characteristics, as well as the circumstances in the households and communities where young adults grew up.

Understanding the determinants of the migration decision at the individual level is empirically challenging in terms of requisite data (Carletto et al. 2007). First, we need information on both migrants and potential migrants in terms of their origins, and their decisions whether or not to migrate. Among those that do move, we want to know their destination and timing of the migration, including the duration spent away from their place of birth and childhood residence. The second requirement is socioeconomic data on migrants (and non-migrants) before their departure. In the best circumstances, panel data is the solution, even though its collection can be very costly (Blumenstock 2012). In the absence of panel data, we exploit a unique module of the 2003 *Education et Bien-être des Ménages au Sénégal* (Education and Household Welfare in Senegal) survey which was specifically designed to understand migration decisions by asking questions retrospectively on young adults (21 to 35 years of age). We specifically rely on data we collected on both migrants and non-migrants when they were 10 years old.

Although most empirical studies and policy debates have focused on rural-to-urban migration, reflecting concerns over the rapid rate of growth of cities, few studies have highlighted the importance of other internal migration patterns such rural-to-rural (Banerjee and Duflo 2007). Likewise, most studies have concentrated on permanent migration, rather than large flows of temporary migrants. In response, we focus on the scope of internal migration and the diverse nature and experience of internal mobility by differentiating between temporary and permanent migrants and whether their origins and destinations were rural or urban. We find that more than half of the total internal migration is temporary and rural to rural or urban to urban, not the more widely studied rural-to-urban permanent migration. We employ a multinomial choice model to empirically model the decision of young people to migrate to either rural or urban areas.

In addition to variables on the individual characteristics, such age, gender, and ethnicity, we include childhood demographic characteristics, such as the number of siblings, the role of the family's financial constraints measured by the asset index of the household when the child was 10 years of age, parents' education, and shocks in terms of death of their father and/or mother. Furthermore, we control for childhood residence characteristics such as access to education and health centers. Empirical evidence in developing countries has shown that women and men may migrate for different reasons (Smith and Thomas 1997; Ritcher and Taylor 2008). Therefore, we estimate separate models for young men and women to test if the socioeconomic determinants are gender-specific.

Our findings reveal that the socioeconomic factors of youth internal migration are heterogeneous by gender. For example, ethnicity and age play differentiated roles for women and men in defining migration choices to either urban or rural areas. Also, fathers' education influences only young women's choices; the higher the fathers' education the more likely are the daughters to move to urban areas, and the less likely they are to move to rural areas. Our results also indicate that family and community circumstances during childhood play a role in predicting migration later in life. For instance, we observe that the number of siblings affects the migration decision:

women with younger siblings are more likely to migrate to rural areas, suggesting that elder sisters can leave home if they have younger siblings who act as substitutes in their social and economic tasks in the household. While the results of these demographic variables are interesting, we emphasize the need for some caution in interpreting these effects causally, as unobserved heterogeneity may jointly affect fertility and migration decisions. We also find that social infrastructure in the community where an individual lived at age of 10, shapes migration decisions; our results show that access to a nearby primary school decreases youth migration to urban areas.

The findings of this paper contribute to the literature that has drawn attention to the importance of household and community characteristics in the individual's decision to migrate, specifically adding to the very few studies that highlight the importance of family and community level characteristics during childhood in predicting later migration in life (Abramitzky et al. 2012). The determinants of youth internal mobility are particularly relevant in Senegal, a country where 65% of the population is less than 25 years old, where mobility is very high, and where poverty and mobility are interrelated. Understanding who is likely to migrate internally, and why, is thus fundamental to formulate education, employment, and other policies to mitigate the stresses associated with such movements and improve the success of new migrants.

The remainder of this paper is organized as follows: In the next section we describe the conceptual framework and empirical approach used in the paper. This is followed in Section 3 by a discussion of the household survey data, including how we define and classify migrants. We also present descriptive data on patterns of internal mobility. The fourth section presents the econometric results from multinomial models that explain the determinants of migration. In the final section of the paper, we conclude and discuss the limitations and implications of this research.

1.2 Conceptual Framework and Empirical Methodology

In contrast to the earlier models of migration that analyze the individual's decision to move as a function of his/her own expected net economic benefit (Harris and Todaro 1970), a growing literature has been modeling migration as both an individual and family decision that not only maximizes income but also minimizes risks (Stark, 1991; Stark and Bloom 1985; Taylor 2001). If migration is an investment decision whereby individuals incur costs to generate higher incomes, youth have lower costs in moving and have higher lifetime expected returns. This is not only because of longer life compared to older people but also because their opportunity cost in the place of origin can be lower due to, for example, high youth unemployment rates. On the other hand, if migration is a family decision and perceived as a risk-coping mechanism, the choice of which household member migrates is based on both earning potential and their ability to be engaged in family insurance arrangements. For instance, Rosenzweig and Stark (1989) show that Indian rural farm households tend to engage in longer distance marriage-cum-migration to cope with volatile profits.

Most family decision migration models thus imply that while a member of the household migrates, the rest of the household which remains in the sending place is focused on maintaining its integrity as an economic and social unit. To explain this type of family decision migration behavior, models typically include variables such as: i) physical and human capital assets, including household demographic composition and information on adults' education, which can be proxies for risk aversion and access to credit; ii) stochastic variables such as weather or agricultural shocks to measure the type of risks that the household copes with; and iii) family networks, included as a form of social capital, that together with the human capital assets can generate differences in the net benefit of the migration decision across households (Ritcher and Taylor 2007). Along the same lines, recent studies have highlighted the role of an individual's childhood family background on their probability of migrating later in life. For instance, Abramitzky et al. (2012), using a novel dataset of the Age of Mass Migration (1850-–1913) from Norway to the US, find evidence that economic and family conditions of an individual's household during childhood, particularly parental wealth and gender composition of siblings, can shape the internal or international migration decisions later in adult life. Following this conceptual framework, we model the young people's decisions to migrate to either rural or urban areas in Senegal as a function of their individual characteristics, the characteristics of their childhood family, and community circumstances prior to their departure.

Empirical studies addressing the determinants of migration face the challenge of observing the individual's migration at one point in time after this decision has been made. Furthermore, migration decisions can be made jointly with other household decisions such as investments in education and resource allocation, raising potential problems of endogeneity between migration and its determinants. We address this issue by using a survey from Senegal that contains retrospective data on young migrants and non-migrants, between 21 and 35 years old. The retrospective information on the household and community, when the individuals were 10 years old, allows us to simulate the impact of an individual's childhood circumstances prior to his/her migration decision.

The decision to migrate and the choice of destination are jointly modeled using a multinomial logit model where individuals can choose between staying (not moving), migrating to a rural area, or migrating to an urban area. More specifically, we estimate the following reduced form regression:

$$Ln\left[\frac{p(M_i = 1,2)}{p(M_i = 0)}\right] = \alpha + \beta' X_i + \delta' E_i + \theta' H_i + \rho' C_i + \pi' R_i + \epsilon_i$$

Where M_i *i*s the destination choice variable such that *i* takes the value of 0 if the individual does not migrate (the base case scenario), 1 if the individual migrates to a rural area, and 2 if the individual migrates to an urban area; X represents individual characteristics; E is parents' education at the time of survey; H and C are other household and community characteristics when the individuals were 10 years old; and R is a set of regional dummies, corresponding to the region of childhood residence, to control for fixed effects that can influence the cost of migration such as weather shocks, and economic and social networks effects. [1]

To amplify, among the individual characteristics we include age, ethnicity, and gender. In the case of the former, we would expect that younger individuals have a

[1] Since the independent variables are from the chooser and not the destination choice, we are not required to implement a test of independence of irrelevant assumptions (IIA).

higher probability of migrating, reflecting, among other considerations, that according to the human capital theory of migration, younger individuals have higher earnings potential. In addition to running the models with a gender dummy variable, we also account for differences in the determinants of migration by estimating separate models for young women and men. In fact, some empirical studies in developing countries have shown that young women, unlike men, frequently move to marry (Smith and Thomas 1997). Also, gender differences are expected when parents encourage daughters, rather than sons, to migrate because of the expectation that the former are more likely to remit (World Bank 2007). The individual's ethnicity is included as a proxy for social networks and cultural norms, which in the case of the Senegal is particularly important as we will describe later. We exclude from the models any current individual's educational attainment because the potential reversal causality with migration.[2]

In terms of household covariates, if migration is considered as a family decision, the education of the father and mother are expected to influence a young person's decision to migrate, and should, therefore, be included among the model determinants (Smith and Thomas 1997; Quisumbing and McNiven 2006). The more educated the parents, the more information is available about the net benefit of migration, which can increase the odds of leaving. Also, parents' education can be a proxy for other household assets such as networks and family connections that can increase the probability of migrating. The empirical evidence is, however, not conclusive on the direction of the parents' education; while Ezra et al. (2001) do not

[2] In the 2003 EMBS survey, we lack a variable to instrument education, or we cannot infer the education completed before the migration decision.

find an effect of the household head's education on Ethiopian rural out-migration, Pessino (1991) shows that the years of education of the mother increases the likelihood to leave rural, but not urban areas in Peru.

To model the childhood environment prior to migration, we include different variables reflecting the household and community characteristics at the time the potential migrants were 10 years old. To measure the wealth and risk aversion of the household, we construct an asset index based on the dwelling conditions at age 10, described in the next section. Our expectation regarding the sign of the wealth effect on the probability of migrating is not unambiguous. On one hand, we can expect that better off households will be less likely to encourage their children to migrate, since the higher their assets, the better the potential economic opportunities in the community in which the young adult resides as a child.[3] On the other hand, we can expect that asset-poor households are less able to finance the costs of migration, and thus, less likely to migrate.[4] Indeed, Mackenzie et al. (2007) show that the probability of migrating from Mexico to the US has an inverse U-shaped relationship with wealth. This nonlinear effect is explained by the heterogeneity of migration networks: in sending communities with lower migration networks, the costs of migrating are relatively high and wealth is positively correlated with the likelihood to migrate, but once the migration networks are larger, the costs, and thus the importance of wealth on the decision to migrate, decrease. We tested in our models for an inverse U-shaped

[3] Other factors that affect this relationship are, for example, the land tenure systems in developing countries that can shape the youth migration decisions. For example, in the Philippines, young adults stay with their parents if they inherit the land (Quisimbuing and McNiven 2006).

[4] Marleto (2008) shows that poorer households in Bangladesh are only able to afford domestic migration while the better off households can afford the costs of international migration.

relationship between the asset index and the probability of migration by introducing a quadratic term in our regressions, but we did not find any statistically significance for this nonlinearity.

Additionally, young adults can migrate as a response to weather risks or any other type of income shock in the sending places (Yan and Choi 2008). Although, we lack information about weather shocks prior to an individual's departure, we do have information on whether either one or both parents had passed away by the time the individual was 10 years old. We capture this by including a dummy variable that takes the value of 1 when the individual reports that their father, mother or both passed away by the time they were 10 years old.[5] We expect that this shock can shape later migration decisions because sons and daughters are likely to assume economic and social responsibilities at home that can trigger or hinder their migration. Some studies in Ghana and Bangladesh show that losing a parent encourages children to move out of their childhood place of residence in search of income-generating activities (MDR 2009).

At the household level, we also include the number and gender composition of the individual's siblings. We do so while acknowledging that the number of siblings can be in part a function of household preferences for the quality and quantity of children. Nonetheless, the question of whether the presence of younger/older male and female siblings contributes to migration provides interesting insights in terms of these relationships, even if we cannot draw strict causal inferences from the results. These results can suggest some underlying mechanisms of intra-household allocation of time

[5] We could not try a separate dummy variable for each parent's death since the number of cases for either mother or father was too small.

and resources, related to, for example, the division of household work and labor market activities, or even marriage practices and cultural norms, that shape an individual's migration decision. For instance, in the context of the migration from Norway to the US in the early nineteenth century, Abramitzky et al. (2012) show that men who had fewer brothers and were the oldest in their families were less likely to migrate later in life; the authors argue that this result is explained by the fact that in Norway the bequest was given to the oldest brother. Younger brothers, having less access to family resources, were more likely to migrate in search of better opportunities. In addition to the household allocation of resources among siblings, there may also be a role played by rights and tasks that relate to a child's birth order position relative to his siblings. For example, Protik and Kuhn (2007) show for Bangladesh that the migration of older brothers decreases the likelihood of sisters to marry and reside in places far from their parents. One explanation they give is that, in order to ensure for elderly care by their daughters, parents might prevent a marriage that involves migration. Furthermore, there might be substitution of tasks among siblings of the same gender that shape migration choices. For example, younger sisters are less likely to migrate since they assume expanded responsibilities for performing household chores when replacing older siblings who have previously migrated (Smith and Thomas 1997; Quisibuing and and McNiven 2006).

At the community level, we include dummy variables on the access to primary and secondary schools and to hospitals, when the young adults were 10 years old. For each one of these variables, access is defined as the existence of the corresponding institution within five kilometers from the individual's residence when he/she was 10 years old. Since public policy determines the geographic distribution and disparity of social infrastructure, these variables are particularly important for understanding the role of government investments in migration choices. For instance, Katz (2000) shows that women in rural areas of Ecuador are more likely to remain in communities with high levels of local organization, such as access to school and hospitals.

1.3 Data Sources and Descriptives

The data we use in this paper is the 2003 Household Survey on Education and Welfare in Senegal (EMBS), conducted in 33 rural and 30 urban communities.[6] Although, as discussed by Glick and Sahn (2009, 2010), the sample is not truly nationally representative since it is part of a cohort study of young children, efforts were made to randomly select into the sample new households to ensure that it is as close as possible to a random sample. Indications from comparison with other national surveys are that this effort was quite successful and that the sample of 1,820 households is representative of the population in terms of religion, ethnic groups, and demographic characteristics, as well as other characteristics such as education.[7]

In our analysis, we rely extensively on the migration module of the EMBS which contains information on the current residence, the birth place, and the residence five years prior the survey (1998). It also provides the years of residence in the current location. In addition, this module has retrospective questions for adults above age 21 about where they lived, as well as the household and community characteristics when

[6] See Glick and Sahn (2007) and (2010) for details about the survey design.

[7] For example, net primary enrollment in our sample (primary enrollments of children 7–12) is 66 percent compared with 63 percent for the country as whole in 2000 (World Bank 2006).

they were 10 years old. This data is a key component of our methodology, because it allows us to observe the childhood characteristics of both migrants and non-migrants that we use to analyze migration decisions.

Defining a migrant in empirical work is not always straightforward and often made difficult due to limitations of the available data. We define *migrant* as an individual who has ever been outside of his/her community for at least one year and departed from their place of origin after he/she was 10 years old.[8] Among our sample of 2,676 individuals who fall in the age group of 21 to 35 years old,[9] 35.01% are defined as *migrants*; in other words, 937 individuals left their communities at least for one year after they were 10 years old; the median age of departure among these young migrants is 20 years.[10]

While most of the empirical studies of internal migration in developing countries have focused on out-migration, especially from rural areas, they have neglected a careful examination of different patterns or types of migration such as rural-to-rural or sequential migration. Mainly, this omission has been justified by the lack of data. Among the few studies in developing countries, Pessino (1991) analyses the determinants of different types of migration in Peru. Identifying the movements by

[8] We test whether our results will change if we exclude the individuals who migrated at younger ages, between 10 and 14 years old. It is plausible that for these individuals, parents might strongly influence or make their decisions to migrate. If this is the case, the migration decision will be endogenous to other household-level decisions, such as fertility. We find that our key results are not sensitive to this choice of including these younger migrants. (See Table A.1.2 in the Appendix.)

[9] We use this age range since previous studies of internal migration have shown that internal flows are the highest for persons in this age group, especially as they search for employment and better economic prospects (Ezra et al. 2001; Brockerhoff et al. 1993); additionally, this cohort is especially important in terms of their experiences and recentness of their moves. We also suspect that the recall data is more accurate for these younger adults than for older individuals.

[10] We calculate the age of departure by subtracting the number of years of residence in the destination (current place) from the young migrant's current age.

the degree of urbanization of the origin, the author finds that primary migrants, i.e., people who move for the first time, are more likely to come from rural areas while repeat or return migrants, i.e., people who have prior moves, come from urban areas. Reed et al. (2010), using a household survey in some regions in Ghana, find that past and future mobility are positively and strongly correlated, suggesting that previous mobility reduces the perceived cost of moving again.

Another important study that attempts to classify migrants is that of Juan and Kim (1979) using census data in the Philippines. The authors construct a comprehensive set of categories of migrants that distinguishes migrants by various characteristics, including the number of moves and whether they return to their birthplace. Building upon this previous work, and using the information from our survey on the place of residence: (1) at the time of the survey (2003), (2) five years prior to the survey (1998), (3) when individuals were 10 years old, and (4) when individuals were born, we first focus on the periodicity of movements—that is, how many times the individual moves across these points in time. We distinguish between primary (one move) and repeat migrants (two or more moves), as well as return migrants. The latter category includes those whose second or third move involved returning to their birthplace. To be included in the category of return migrants, by definition, they have to report having lived at a location other than their birthplace either when they were 10 year of age, in 1998, and/or at the time of the survey. In our sample, 25.4% are primary migrants, 3.0% are secondary or tertiary migrants, and 11.9% are return migrants. A final, and the largest group of migrants—fully 59.6%—are those who we define as "temporary return" migrants, but for whom we do not have information on their migration other than they were away from their birthplace for at least one year. Thus, these individuals report that they were both resident in another location for at least one year, but also that their birthplace is the same as their residence at the time of the survey, and that they lived in their birthplace in 1998 as well as when they were 10 years old. [11]

Table 1.1 shows the distribution of migration by the urban/rural origin and destination of the move, as well as the migration categories: primary, return, repeat and temporary, discussed above. We find that two-thirds of the migrants from rural to urban areas are primary migrants; this is consistent with the fact that most of the migrants in Dakar are more likely permanent migrants (World Bank 2006). On the other hand, almost 60% of the rural-to-rural and urban-to-urban migrants are temporary movers.

Table 1.1 Distribution of migrants by rural/urban birthplace and urban/rural residence in 2003

	Primary	Repeat	Return	Temporary	Total
Urban -Rural	60.6%	3.0%	1.5%	34.8%	100.0%
Rural- Rural	7.3%	0.5%	14.7%	77.5%	100.0%
Urban -Urban	26.2%	4.3%	14.5%	55.1%	100.0%
Rural -urban	59.8%	8.3%	2.3%	29.5%	100.0%

Source: Authors calculations based on 2003 EMBS

Table 1.2 summarizes some socio-economic characteristics of migrants and non-migrants among our 21- and 35-year-old cohort of young adults. We include in

[11] Juan and Kim (1979) (as explained in Blisborrow (1984)) classify these persons as non-migrants, because they report the same place of residence at all points of time that are included in the survey. We acknowledge that there may be some misreporting among this group —that is, that they made an error in reporting having lived elsewhere for more than one year. However, we expect that the vast majority answered that question correctly, and are indeed return migrants who happened not to live away from their place of birth in 1998 and when they were 10 years of age. In our analysis, we explore whether the results are sensitive to the inclusion/exclusion of these groups being characterized as migrants.

these tables, and in the analysis that follows, temporary migrants, having tested that the results would not be appreciably altered by excluding the temporary migrants under the assumption that this group might have different triggers to migrate internally. (See Table A.1.3 in the Appendix.) Our young migrants are mostly female, although they have similar ethnic distribution and education levels compared to the non-migrants.

Table 1.2 Socioeconomic characteristics of migrants and nonmigrants

	Migrants	**Nonmigrants**	**Total**
Characteristics in 2003- year of the survey			
Percentage of female	64.78	53.36	57.36
Average age	27.79	26.39	26.88
Average years of education	4.15	4.50	4.38
Ethnicity			
% belong to Wolof	29.39	35.77	33.47
% belong to Poular	24.71	19.98	21.68
% belong to Mandingue	17.84	13.40	15.00
% belong to Serere	16.23	20.39	18.89
Parents Education			
% whose mother has none school	85.37	82.17	83.28
% whose father has none school	72.91	68.30	69.90
Access to Social services at age of 10 years			
Access to Primary School	0.86	0.95	0.91
Access to lower Secondary	0.46	0.55	0.51
Access to Upper School	0.35	0.46	0.42
Access to Hospital	0.72	0.84	0.79
Distribution by asset quartiles at age 10 years			
(percentage of Individuals)			
First	40.77	30.76	34.27
Second	18.36	17.42	17.75
Third	19.10	27.26	24.40
Fourth	21.77	24.55	23.58

Source: Authors' calculation from 2003 EMBS

Women represent more than two-thirds of the young migrants, compared to 53.3% in the non- migrant group and 57.3% in the total population. This female overrepresentation in the group of young migrants can be presumably explained by the association of migration and the decision to marry, as we will discuss further in the next section.

Migrants are slightly older than non-migrants in our sample. However, this difference is not statistically significant. In addition, there is no age difference between men and women within the migrant and non-migrant groups. Among the migrants, almost 30% are Wolof, 24 % Poular/Toucolder, 17% Mandingue, and 17% Serere, with the remainder belonging to other minority groups. This distribution is very similar for the non-migrants and the total sample. Some ethnographic evidence has shown an association between ethnicity and migration in West Africa. On one hand, Bockefort et al. (1993) document that young Serere and Diola women are more prone to migrate, especially seasonally, to urban areas for domestic work, while women belonging to Toucoluer and Sonike groups are less likely to move from the village. Similarly, the authors show that women belonging to Wolof groups are less prone to migrate. We empirically investigate this further in our econometric models.

We also see that non-migrants have slightly more education; however, it is not statistically significantly different from migrants. More than 70% of the migrants' fathers did not go to school, and this percentage is even higher, 85%, for their mothers. This situation is not appreciably different for non-migrant young adults; 68% of their fathers and 83% of their mothers did not go to school.

Descriptive statistics on the access to social infrastructure when young migrants and non-migrants were 10 years old indicate that migrants come from areas with less access to a nearby primary school, lower secondary school, and upper secondary school, as well as a hospital. Approximately 91% of the young people had a primary school near their residence. However, this percentage is only 86% for the migrant youth. Along the same lines, 46% of young migrants came from a region with a middle school (lower secondary) nearby while this percentage was almost 10 points higher for the non-migrants. Similar patterns are found when analyzing access to upper secondary school. The access to health services was also unequal between migrants and non-migrants in their childhood residences. While 72% of the migrant had access to a hospital, this percentage was 84% for the non- migrant population.

As noted above, we create an asset index following standard procedures using factor analysis and the dwelling characteristics where the young adults lived at 10 years of age.[12] While 40% of the migrant children came from the lowest quartile, this percentage was 30% among the non-migrant group. However, this difference seems to be smaller for the highest quartile. Overall, we find that the non-migrant's asset distribution first order dominates the migrant's.

1.4 Results and Discussion

The results of the multinomial logit models are presented in Appendix Table A.1.1, with the marginal effects shown in Table 1.3. Panel A shows the marginal effects for all the individuals between 21 and 35 years old; panels B and C show the results for young men and women, respectively. Given that the marginal effects are more meaningful and easier to interpret, we will concentrate our discussion on these estimates.

[12] We construct the asset index based on the floor material, the source of potable water, and the type of bathroom for the dwelling. These were the only characteristics available in the retrospective survey module.

Table 1.3 Multinomial logit models by rural and urban destinations-Marginal Effects

	Panel a		Panel b		Panel c	
	All		Men		Women	
	Rural	Urban	Rural	Urban	Rural	Urban
Individual Characteristics						
Gender (1=male)	-0.064***	-0.013				
	(0.014)[a]	(0.015)				
Age	0.076***	-0.052**	0.041	-0.025	0.099***	-0.063**
	(0.022)	(0.024)	(0.025)	(0.035)	(0.034)	(0.031)
Age squared	-0.001***	0.001**	-0.001	0.001	-0.002***	0.001**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Wolof	-0.016	-0.062*	-0.018	-0.122***	-0.020	-0.012
	(0.030)	(0.032)	(0.035)	(0.044)	(0.046)	(0.044)
Poular	-0.019	-0.007	0.033	-0.038	-0.087**	0.025
	(0.025)	(0.032)	(0.026)	(0.044)	(0.040)	(0.045)
Serere	0.007	-0.168***	-0.023	-0.155***	0.038	-0.174***
	(0.032)	(0.042)	(0.036)	(0.056)	(0.051)	(0.059)
Diola	-0.024	-0.004	0.014	-0.015	-0.055	0.004
	(0.038)	(0.044)	(0.041)	(0.061)	(0.060)	(0.062)
Other	-0.092*	-0.080*	-0.057	-0.145**	-0.132*	-0.016
	(0.048)	(0.048)	(0.062)	(0.070)	(0.071)	(0.064)
Missing ethnicity	0.004	-0.074**	0.021	-0.118***	-0.014	-0.034
	(0.026)	(0.031)	(0.029)	(0.042)	(0.042)	(0.043)
Household Characteristics						
Father's highest education	-0.017***	0.011***	-0.013*	0.005	-0.022**	0.017***
	(0.006)	(0.004)	(0.007)	(0.006)	(0.009)	(0.005)
Mother's highest education	-0.005	0.005	0.004	-0.006	-0.011	0.010
	(0.009)	(0.006)	(0.010)	(0.008)	(0.014)	(0.008)
Asset Index	-0.008	0.028**	-0.018	0.028	-0.002	0.022
	(0.011)	(0.012)	(0.014)	(0.018)	(0.017)	(0.016)
Being Orphan	0.065***	0.038	0.052	0.068	0.080**	0.020
	(0.025)	(0.031)	(0.032)	(0.049)	(0.037)	(0.039)
No. older siblings	-0.002	-0.002	-0.007*	-0.006	0.004	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.005)
No. younger siblings	0.007**	0.001	0.001	0.005	0.012**	-0.003
	(0.003)	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)

Table 1.3 (continued)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Community Characteristics						
Access-Primary School (1=yes)	-0.023	-0.178***	0.020	-0.197***	-0.064*	-0.173***
	(0.022)	(0.036)	(0.026)	(0.056)	(0.034)	(0.046)
Access-Secondary School(1=yes)	0.006	0.002	0.034	-0.073*	-0.025	0.074*
	(0.023)	(0.029)	(0.025)	(0.039)	(0.037)	(0.041)
Access-Hospital(1=yes)	-0.069***	0.046	-0.031	0.088*	-0.099***	0.021
	(0.018)	(0.034)	(0.020)	(0.052)	(0.029)	(0.045)
Rural residence 10yrs (1=yes)	0.148***	-0.097***	0.130***	-0.142***	0.139***	-0.055
	(0.027)	(0.029)	(0.032)	(0.041)	(0.042)	(0.041)
Regions						
Diourbel	-0.013	0.054	-0.045	0.044	0.003	0.055
	(0.038)	(0.034)	(0.051)	(0.050)	(0.056)	(0.046)
Fatick	-0.014	0.161***	-0.026	0.098	-0.003	0.201***
	(0.046)	(0.044)	(0.060)	(0.065)	(0.070)	(0.060)
Kaolack	0.006	0.033	0.008	-0.009	-0.001	0.061*
	(0.034)	(0.026)	(0.040)	(0.038)	(0.051)	(0.035)
Kolda	0.101***	-0.055	0.035	-0.061	0.157***	-0.059
	(0.035)	(0.035)	(0.043)	(0.051)	(0.053)	(0.048)
Louga/Matam	-0.014	0.008	-0.043	0.020	0.015	0.002
	(0.039)	(0.033)	(0.053)	(0.046)	(0.056)	(0.045)
Saint Louise	-0.031	0.030	-0.027	0.035	-0.021	0.027
	(0.040)	(0.030)	(0.049)	(0.044)	(0.059)	(0.041)
Tambacounda	0.023	-0.127***	0.032	-0.169**	-0.000	-0.097*
	(0.038)	(0.045)	(0.045)	(0.067)	(0.060)	(0.058)
Thies	0.019	0.003	0.030	-0.028	0.016	0.021
	(0.032)	(0.026)	(0.040)	(0.041)	(0.047)	(0.034)
Ziguinchor	0.125***	0.144***	0.088*	0.133**	0.129*	0.157***
	(0.042)	(0.037)	(0.048)	(0.055)	(0.067)	(0.049)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively.
[a] Standard errors of marginal effects are calculated by the delta method and reported in parentheses
No observations: All, 2429; Men: 1047; Women: 1382

1.4.1 Individual Characteristics

From the model that includes both men and women, the negative and significant gender variable indicates that women are 6.4% more likely than men to move to rural areas, although no gender difference exists for moves to urban areas. These results may reflect that young women often move as a consequence of following their spouses. While we are unable to prove the causal effect of marriage on female youth migration, previous research indicates that marriage is the main reason for migration among women between 15 and 49 years old in Senegal (Safir 2009). We examined the relationship between the age of marriage and age of migration. First, we note that on average, among married couples, men are 12 years older.[13] Second, we notice that 72 percent of the women who migrate were already married, in contrast to only 31 percent of the male migrants.

We also examine the marginal effect of age among the cohort of individuals between 21 and 35 years old, as shown in Panel A, being one year older increases by 7.6% the probability of migrating to rural areas and decreases by 5% the probability of migrating to urban areas. While age has no effect for men on the likelihood to migrate to either rural or urban areas, for women this effect varies with their destination. As age increases, women are 10% more likely to migrate to rural areas and 6% less likely to migrate to urban areas; however, this effect is non-linear as seen by the statistical significance of the quadratic term.

[13] In the 2003 EMBS sample of married couples, the average woman's age is 38 while for men, it is 50 years old.

The results also show evidence that ethnicity influences the likelihood of migrating to rural and urban areas[14]. This effect is differentiated by gender. On one hand, belonging to the Serere group, relative to the Mendingue/Sose group that was excluded, decreases by 17% the likelihood of migrating to urban areas. This marginal effect has similar magnitudes among women and men. On the other hand, belonging to the Wolof group decreases only male migration to urban areas by 12%, while belonging to the Poular group decreases only female migration to rural areas by 9%. Also, men who belong to other ethnic minority groups have a 15% lower probability of moving to urban areas while women belonging to the same groups have a 13% lower probability to go to rural areas. These results are consistent with the study by Bockefort et al. (1993) who discuss the importance of ethnicity in the decision for women to migrate in Sub-Saharan countries, including Senegal.

1.4.2 Demographic and Economic Household Characteristics

Our results indicate that the children of fathers with more education are less likely to move to rural areas, and more likely to move to urban areas, where the magnitude of these effects are very similar. Mother's education, however, is not significant[15]. When examining the gender-disaggregated results, we observe that the effect of the father's education on youth migration is larger and more statistically robust for their daughters

[14] We also include a dummy variable for missing observations given the substantial amount of misreporting of this variable in the sample (523 observations for non-migrants and 253 for migrants).

[15] We corroborate these results by estimating the same multinomial model (not shown) but including a dummy variable for each father and mother if he/she is literate (i.e., has some level of education).

than it is for their sons.[16] This result may reflect the role of fathers in arranged marriages, or perhaps in terms of promoting more educational opportunities for their daughters, which often requires migrating to urban areas. These two mechanisms, in fact, may be related: greater education of the fathers, whether it be through ability, economic well-being, or more expansive social networks, may enable them to find more favorable husbands for their daughters who will move with their husbands to the city in pursuit of greater opportunity, or similarly, to improve educational opportunities for their daughters, which requires schooling in urban areas. In contrast, fathers' education may discourage marriage arrangements whereby daughters migrate to rural areas where the returns of migration areas are likely to be lower. These results are similar to those found by Quisimbuing et al. (2005) in the Philippines where fathers' education increases the probability of daughters moving from the village, and interestingly, mothers' education had the opposite effect.

Our models also suggest that better living conditions during childhood, measured by the dwelling asset index, are associated with higher likelihood of migrating to urban areas, while decreasing the likelihood of migrating to rural areas; however, the latter effect is not statistically significant[17]. The asset index does not have a differentiated effect by gender. This result might suggest that young women and men who grew up in asset-poor households are less able to afford the costs of migration to urban areas. We also test if there was a differentiated effect of the asset

[16] The effect of the father's education on young males is significant only at 10%, and it is not robust to the specification of a father's literacy dummy variable.

[17]This result is consistent with the fact that the asset distribution for migrants going to urban areas first order dominates the migrants going to rural areas.

index by rural/urban origin. A better-off asset position of the household in the rural origin decreases the likelihood of migrating to either rural or urban areas. Interestingly, this effect is statistically significant for men and not for women, suggesting that male migration might be deterred by better economic opportunities, probably associated with agricultural activities, in the rural areas. (See Table A.1.4 in the Appendix.)

The multinomial regressions shown in Table 1.3 also include information on the demographic make-up of the households when individuals were 10 years old. The results indicate that a higher number of younger siblings increase the probability of migrating to rural areas, while a higher number of older siblings does not have any effect on the probability of migrating to either urban or rural areas. Looking at the models by gender, the results show that the marginal effect of having younger siblings is still statistically significant for women; and this effect is only positive and significant in the case of women moving to rural areas. Also, the presence of older siblings only decreases male migration to urban areas.[18]

One possible explanation is that women with a higher number of younger sisters are more likely to migrate because their young female siblings act as substitutes in home production (Smith and Thomas 1997; Quisminbuing 2005). Indeed, we further examine the sex and birth order composition of the siblings in the likelihood of migration. We estimate the multinomial models including younger and older brothers and sisters. See Table 1.4. We find that having younger sisters increases the odds of moving to rural areas, and this effect is significant for women, but not for men.

[18] However, this result is not statistically robust when disaggregating the siblings by gender and birth order composition.

Table 1. 4 Multinomial logit models with siblings' composition (Marginal effects)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Household Characteristics						
Father's highest	-0.017***	0.011***	-0.013**	0.005	-0.024**	0.017***
education	(0.006)[a]	(0.004)	(0.006)	(0.006)	(0.009)	(0.005)
Mother's highest	-0.004	0.004	0.004	-0.006	-0.011	0.010
education	(0.009)	(0.006)	(0.010)	(0.008)	(0.014)	(0.008)
Asset Index	-0.008	0.028**	-0.017	0.029*	-0.002	0.022
	(0.011)	(0.012)	(0.014)	(0.018)	(0.017)	(0.016)
Being Orphan	0.066***	0.040	0.048	0.069	0.081**	0.021
	(0.025)	(0.031)	(0.032)	(0.049)	(0.037)	(0.039)
No. of older	0.002	0.006	-0.012	-0.000	0.016*	0.012
brothers	(0.006)	(0.006)	(0.007)	(0.008)	(0.009)	(0.008)
No. of older	-0.006	-0.012*	-0.002	-0.014	-0.010	-0.013
sisters	(0.007)	(0.007)	(0.008)	(0.010)	(0.010)	(0.009)
No. of younger	0.004	-0.001	-0.001	-0.000	0.010	-0.005
Brothers	(0.005)	(0.006)	(0.005)	(0.007)	(0.008)	(0.008)
No. of younger	0.011**	0.002	0.004	0.010	0.015*	-0.003
sisters	(0.005)	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)

***, **, *: significant at 1%, 5%, 10% levels. [a] Standard errors of marginal effects calculated by the delta method and reported in parentheses. Marginal effects of individual and communities characteristics as well as regional dummies are not shown. N° observations: All, 2429; Men: 1047; Women: 1382; Pseudo R^2: All, 0.1497; Men, 0.1598; Women, 0.1594.

In addition, we account for whether the individual is an orphan of either father or mother, or both, by the time s/he was 10 years old. Our results indicate that the marginal effect of the loss of a parent during childhood increases by 6.3% the probability of migrating to rural areas, but it does not affect the likelihood of moving to urban areas. By gender, we find that being an orphan only affects female and not male migration, and this effect is only significant for those women going to rural areas.

1.4.3 Community Characteristics

The availability of social infrastructure, such as schools and hospitals, in the community where the individual lived as a child influences their probability of moving. Access to a primary school within five kilometers decreases the likelihood of moving to urban areas by 18%, but it does not affect the probability of moving to rural areas. This marginal effect is of a similar magnitude for both men and women, although it is no longer statistically significant for men migrating to rural areas. We also investigate whether the nearest primary school has a differentiated effect on the likelihood to migrate by whether the individual lived in a rural or urban area as a child. To do so we estimate the models (not shown) including an interaction between the urban dummy and the nearest primary school. We find that proximity to primary school decreases the probability of migrating to urban areas only if the early childhood residence is in a rural area.

Access to secondary school does not affect the decision to migrate in the aggregate sample; however, when we examine the gender-disaggregated models we see that a secondary school within five kilometers actually increases the female probability of migration to urban areas by 7%, and by a similar magnitude, decreases the male probability of migration to the same areas. We expect that this is mediated by the fact that access to secondary schools exposes girls and their families to the potential of greater opportunities associated with education and a greater openness to migrate in search of opportunity, whether in the labor market or in search of more education. Proximity to a nearby hospital decreases the odds of migrating to rural areas only, but again, this is only the case for potential women migrants.

In general terms, our results indicate that better access to social infrastructure during childhood, particularly primary schools and hospitals, deters later youth migration, a finding which is consistent with other empirical evidence in developing countries (Katz 2000). However, there are potential countervailing forces that might contribute to better social infrastructure encouraging migration: that is, easier access to schools can also trigger migration if individuals who accumulate more human capital in presence of schools migrate to other places to look for higher returns to their capital accumulation. In fact, we find that women with access to secondary school when they are 10 years of age are likely to migrate to urban areas.

The models in Table 1.3 also show that the region where the individuals were living at age 10 plays a role in shaping their decision to migrate since it can reflect differences in economic opportunities and social networks that directly affect the cost and benefits of migration. For instance, our results suggest that the probability to migrate to urban areas increases by 15% (12% for rural areas) for young people who lived in Ziguinichor. This region is located in Casamence, an area where poverty is persistent and conflict has been present for the last 30 years (World Bank 2006). Finally, the dummy for whether the childhood residence was either rural or urban corroborates the patterns described earlier: when the childhood residence is rural, the likelihood to migrate to other rural areas increases by 15%; similarly, when the residence is urban, the likelihood to migrate to urban areas increases by 10%.

1.5 Conclusions

Our goal in this paper has been to highlight the importance and magnitude of internal migration in Senegal, and to analyze the socio-economic determinants of the decisions of young people to internally migrate. Young women, in particular, are far more likely to migrate internally, and the determinants of migration also differ by gender. We focus on the role of household and community characteristics prior to the decision to migrate, using household survey data from Senegal which includes retrospective information from when individuals were 10 years old. Our multinomial logit model allows for individuals, between 21 and 35 years old, to decide between not migrating, and moving to rural and urban areas in Senegal.

Our findings suggest that the decision to migrate internally in Senegal is heterogeneous by gender, and that the determinants of migration differ for those leaving their childhood residence for an urban or rural destination. For example, father's education has an important role in women's migration choices. The more educated the father, the more (less) likely are the daughters to go to urban (rural) areas. In the context of our sample, where 72 percent of the female migrants are married, this result could suggest that a father's education is influential not only in marriage arrangements, but also in the probability that his daughter will marry someone and leave the childhood residence with their new husband in search of greater economic opportunity in urban areas. However, this is only conjecture, as we do not have further information to disentangle the role of marriage and economic opportunity in the decision to migrate. Young people who are orphans are also more likely to migrate, presumably reflecting weaker ties to their childhood places of residence.

The characteristics of the community in which children reside shape migration decisions. Proximity to better social infrastructure during childhood, particularly primary schools and hospitals, are generally associated with a lower probability of migrating. The one clear exception is access to secondary schools, which in fact increases the probability of migration to urban areas for young women. While, on the one hand, proximity to secondary schools may mitigate the need to migrate in search of more education, such accessibility is likely associated with higher schooling attainment, especially for girls whose parents are more reluctant to send their daughters away to boarding schools and/or reside with relatives in order to improve school attainment. These human capital investments may subsequently encourage migration of the young women to urban areas in search of employment opportunities that exploit their human capital investments in education.

Our findings suggest that the presence of younger siblings during childhood is associated with migration decisions. For instance, women with younger sisters (but not brothers) are more likely to migrate, suggesting that younger female siblings act as substitutes in household responsibilities. We also find that those who lived in households with a higher asset index are more likely to migrate to urban areas, possibly because these young women and men are able to finance the costs of migrating to urban areas and to reap the benefits of better opportunities in these areas. Nevertheless, controlling for educational background and infrastructure where a child grows up, there is less of a push factor in the migration decision for those households with better living conditions in rural areas, especially for men.

While these is still much to be learned about the internal migration of young people in Senegal, and more generally in other developing countries, the high degree of mobility and the recognition of certain factors that contribute to these population movements is important for policymakers, both in terms of affecting and planning for the widespread migration. While there remain many questions about the determinants of migration, and how to cope with the stresses on communities and households affected by these population movements, there is every reason to expect that they will only accelerate in years to come.

REFERENCES

Abramitzky, R., Boustan, L. P., & Eriksson, K. (2013). Have the poor always been less likely to migrate? Evidence from inheritance practices during the Age of Mass Migration. *Journal of Development Economics*, *102*, 2-14.

Bilsborrow, R. E., Oberai, A. S., & Standing, G. (1984). Migration surveys in low income countries: guidelines for survey and questionnaire design

Blumenstock, J. E. (2012). Inferring patterns of internal migration from mobile phone call records: Evidence from Rwanda. *Information Technology for Development*,*18*(2), 107-125.

Banerjee, A. V., & Duflo, E. (2007). The economic lives of the poor. *The journal of economic perspectives: a journal of the American Economic Association*, *21*(1), 141.

Brockerhoff, M., & Eu, H. (1993). Demographic and socioeconomic determinants of female rural to urban migration in Sub-Saharan Africa. *International Migration Review*, 557-577.

Carletto, G., and De Brauw, A. (2007) " Measuring Migration Using Household Surveys." *Migration Operational Note 2*. World Bank,Washington, DC:

Deshingkar, Priya, and Sven. Grim. 2004. "Voluntary Internal Migration: An Update. Working Paper, Overseas Development Institute, London.

Development Research Centre on Migration, Globalisation and Poverty. 2009. *Making Migration Work for Development*. Available at: http://www.migrationdrc.org/publications/misc/Making_Migration_Work_for_Development.pdf

Ezra, M., & Kiros, G. E. (2001). Rural Out-migration in the Drought Prone Areas of Ethiopia: A Multilevel Analysis1. *International Migration Review*, *35*(3), 749-771.

Giles, J., & Mu, R. (2007). Elderly parent health and the migration decisions of adult children: Evidence from rural China. *Demography*, *44*(2), 265-288.

Harris, J. R., & Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American Economic Review*, 126-142.

Juan, M. and Kim, J.(1970). *Migration Differentials by Migration Types in the Philippines: A Study of Migration Typology and Differentials*. UNFPA-NCSO population research project, United Nations Fund for Population Activities, The Philippines.

Katz, E. (2000). Individual, Household and Community-Level Determinants of Migration in Ecuador: Are there Gender Differences?. In *Annual Meeting of the Population Association of America. March* (pp. 23-25)

Mendola, M. (2008). Migration and technological change in rural households: Complements or substitutes?. *Journal of Development Economics*, *85*(1), 150-175.

McKenzie, D., & Rapoport, H. (2007). Network effects and the dynamics of migration and inequality: theory and evidence from Mexico. *Journal of Development Economics*, *84*(1), 1-24.

Pessino, C. (1991). Sequential migration theory and evidence from Peru. *Journal of Development Economics*, *36*(1), 55-87.

Protik, A., & Kuhn, R. (2006). The hidden cost of migration: Effect of brother's migration on sister's marriage outcomes in rural Bangladesh. *Unpublished manuscript, Department of Economics, Brown University*.

Quisumbing, A. and McNiven, S. (2006). Migration and the Rural-Urban Continuum: Evidence from Bukidnon, Philippines. *Philippine Journal of Development* (2):1–43.

Reed, E., Andrzejewski,C and White,M. (2005). "An Event History Analysis of Internal Migration in Ghana: Determinants of Interregional Mobility among Residents of Coastal Central Region." International Union for the Scientific Study of Population (IUSSP), XXVth Population Conference, Tours, France, 2005. Available at: http://iussp2005.princeton.edu/papers/52292

Rosenzweig, M. R., & Stark, O. (1989). Consumption smoothing, migration, and marriage: Evidence from rural India. *The Journal of Political Economy*, 905-926.

Pfeiffer, L. and Taylor,E. (2008). "Gender and the Impacts of International Migration: Evidence from Rural Mexico." in Andrew R. Morrison, Maurice Schiff, Mirja Sjöblom (eds.), *The International Migration of Women*, New York: Palgrave McMillan: 99-124.

Safir, A. (2009) Who leaves, who moves in? The impact of positive and negative income shocks on migration in Senegal. World Bank Working Paper No 76, Washington, DC.

Sahn, D. E., & Stifel, D. (2003).Exploring alternative measures of welfare in the absence of expenditure data. *Review of income and wealth*, *49*(4), 463-489.

Smith, J. P., & Thomas, D. (1998). On the road: marriage and mobility in Malaysia. *Journal of Human Resources*, 805-832.

Stark, O., & Bloom, D. E. (1985). The new economics of labor migration. *The American Economic Review*, 173-178.

Stark, O. (1991). *The migration of labor* (pp. 216-220). Oxford: Blackwell.

Taylor, J. E. (1999).The New Economics of Labor Migration and the Role of Remittances in the Migration Process. *International Migration* 37(1): 63–88.

__________. (2001) "Migration: New Dimensions and Characteristics, Causes, Consequences and Implications for Rural Poverty." in *Food, Agriculture, and Rural Development: Current and Emerging Issues for Economic Analysis and Policy Research,* edited by Kostas G. Stamoulis. Rome: Food and Agriculture Organization of the United Nations (Food, Agriculture and Rural Development). Available at: http://www.fao.org/docrep/003/X9808e/x9808e07.htm#ee

United Nations Development Program –UNDP(2009). Overcoming Barriers: Human Mobility and Development. Human Development Report. New York: Palgrave Macmillan.

Walker, J. (2005) .Internal Migration. Working Paper No.5307, Department of Economics, University of Wisconsin.

World Bank (2006). Senegal Second Poverty Reduction Paper Strategy. World Bank Publication, Washington, DC.

__________ (2007). Development and the Next Generation. World Development Report, World Bank Publication. Washington, DC

__________ (2010) Senegal Country Report. World Bank Publication. Washington, DC.

APPENDIX

Table A.1.1 Coefficients of Multinomial logit models by rural and urban destinations

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Individual Characteristics						
Gender (1=male)	-0.617***	-0.192				
	(0.13) [a]	(0.12)				
Age	0.639***	-0.311	0.543	-0.165	0.657***	-0.358
	(0.20)	(0.19)	(0.35)	(0.30)	(0.25)	(0.25)
Age squared	-0.010***	0.007*	-0.008	0.004	-0.011**	0.007*
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)
Wolof	-0.235	-0.520**	-0.412	-1.073***	-0.172	-0.129
	(0.27)	(0.25)	(0.50)	(0.38)	(0.34)	(0.35)
Poular	-0.182	-0.080	0.416	-0.284	-0.622**	0.076
	(0.23)	(0.26)	(0.36)	(0.38)	(0.30)	(0.35)
Serere	-0.171	-1.345***	-0.526	-1.372***	0.035	-1.339***
	(0.30)	(0.34)	(0.51)	(0.49)	(0.38)	(0.47)
Diola	-0.231	-0.066	0.177	-0.115	-0.412	-0.045
	(0.35)	(0.36)	(0.59)	(0.53)	(0.45)	(0.49)
Other	-0.971**	-0.775**	-0.996	-1.322**	-1.021*	-0.319
	(0.44)	(0.38)	(0.87)	(0.60)	(0.53)	(0.50)
Missing ethnicity	-0.062	-0.587**	0.147	-0.995***	-0.153	-0.290
	(0.24)	(0.24)	(0.42)	(0.36)	(0.31)	(0.34)
Household Characteristics						
Father's highest education	-0.139**	0.066**	-0.172*	0.025	-0.145**	0.104***
	(0.06)	(0.03)	(0.09)	(0.05)	(0.07)	(0.04)
Mother's highest education	-0.037	0.030	0.050	-0.047	-0.069	0.063
	(0.08)	(0.04)	(0.14)	(0.07)	(0.11)	(0.06)
Asset Index	-0.034	0.216**	-0.213	0.223	0.020	0.176
	(0.10)	(0.10)	(0.20)	(0.15)	(0.13)	(0.13)
Being Orphan	0.662***	0.393	0.819*	0.655	0.637**	0.280
	(0.23)	(0.25)	(0.45)	(0.42)	(0.27)	(0.31)
No. older siblings	-0.019	-0.019	-0.112*	-0.062	0.030	0.011
	(0.03)	(0.03)	(0.06)	(0.05)	(0.04)	(0.04)
No. of younger siblings	0.069**	0.017	0.025	0.042	0.089**	-0.009
	(0.03)	(0.03)	(0.05)	(0.04)	(0.04)	(0.04)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors reported in parentheses. N° observations: All, 2429; Men: 1047; Women: 1382; Pseudo R^2: All, 0.1482; Men, 0.1584; Women, 0.156

Table A.1.1 (Continued)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Community Characteristics						
Access-Primary School (1=yes)	-0.457**	-1.469***	0.033	-1.675***	-0.734***	-1.484***
	(0.20) [a]	(0.30)	(0.36)	(0.51)	(0.26)	(0.38)
Access-Secondary School(1=yes)	0.063	0.025	0.387	-0.588*	-0.081	0.554*
	(0.21)	(0.23)	(0.36)	(0.33)	(0.27)	(0.32)
Access- Hospital (1=yes)	-0.580***	0.276	-0.329	0.717	-0.721***	0.025
	(0.17)	(0.27)	(0.28)	(0.46)	(0.21)	(0.35)
Rural residence 10yrs (1=yes)	1.243***	-0.575**	1.663***	-1.066***	0.975***	-0.238
	(0.26)	(0.24)	(0.46)	(0.36)	(0.32)	(0.32)
Regions						
Diourbel	-0.051	0.419	-0.589	0.324	0.106	0.449
	(0.35)	(0.27)	(0.71)	(0.43)	(0.41)	(0.36)
Fatick	0.096	1.278***	-0.245	0.815	0.272	1.609***
	(0.43)	(0.36)	(0.84)	(0.56)	(0.53)	(0.49)
Kaolack	0.102	0.272	0.105	-0.064	0.082	0.486*
	(0.31)	(0.21)	(0.57)	(0.33)	(0.38)	(0.27)
Kolda	0.864***	-0.305	0.414	-0.480	1.101***	-0.242
	(0.33)	(0.28)	(0.60)	(0.44)	(0.39)	(0.38)
Louga/Matam	-0.122	0.047	-0.590	0.122	0.114	0.039
	(0.36)	(0.26)	(0.74)	(0.40)	(0.42)	(0.35)
Saint Louise	-0.248	0.199	-0.337	0.266	-0.118	0.183
	(0.37)	(0.24)	(0.69)	(0.37)	(0.44)	(0.32)
Tambacounda	0.038	-0.995***	0.247	-1.416**	-0.142	-0.778*
	(0.35)	(0.36)	(0.63)	(0.59)	(0.45)	(0.46)
Thies	0.182	0.050	0.393	-0.200	0.150	0.191
	(0.29)	(0.21)	(0.56)	(0.35)	(0.35)	(0.27)
Ziguinchor	1.366***	1.336***	1.426**	1.265***	1.202**	1.444***
	(0.39)	(0.30)	(0.69)	(0.47)	(0.50)	(0.39)
Constant	-10.953***	3.686	-11.820**	2.335	-10.446***	3.685
	(2.85)	(2.62)	(4.95)	(4.15)	(3.57)	(3.46)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors reported in parentheses. N° observations: All, 2429; Men: 1047; Women: 1382; Pseudo R^2: All, 0.1482; Men, 0.1584; Women, 0.156

Table A.1.2 Multinomial logit models excluding younger individuals (Marginal effects)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Individual Characteristics						
Gender (1=male)	-0.061***	-0.017				
	(0.013) [a]	(0.014)				
Age	0.066***	-0.021	0.035*	0.018	0.086**	-0.041
	(0.021)	(0.023)	(0.021)	(0.033)	(0.034)	(0.031)
Age squared	-0.001***	0.001	-0.000	-0.000	-0.001**	0.001
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Wolof	-0.027	-0.064**	-0.020	-0.127***	-0.041	-0.006
	(0.029)	(0.031)	(0.031)	(0.041)	(0.046)	(0.045)
Poular	-0.023	-0.012	0.030	-0.049	-0.105***	0.031
	(0.023)	(0.031)	(0.021)	(0.040)	(0.040)	(0.045)
Serere	-0.009	-0.138***	-0.015	-0.143***	0.005	-0.123**
	(0.031)	(0.039)	(0.031)	(0.050)	(0.050)	(0.058)
Diola	-0.036	0.002	0.013	-0.026	-0.078	0.032
	(0.036)	(0.042)	(0.033)	(0.056)	(0.059)	(0.061)
Other	-0.120**	-0.083*	-0.031	-0.163**	-0.216**	0.001
	(0.052)	(0.047)	(0.050)	(0.070)	(0.086)	(0.063)
Missing ethnicity	-0.004	-0.071**	0.021	-0.110***	-0.033	-0.027
	(0.025)	(0.030)	(0.025)	(0.038)	(0.042)	(0.044)
Household Characteristics						
Father's highest	-0.012**	0.009**	-0.008	0.004	-0.017**	0.013***
education	(0.006)	(0.004)	(0.006)	(0.005)	(0.009)	(0.005)
Mother's highest	-0.005	0.004	-0.002	0.000	-0.005	0.006
education	(0.009)	(0.005)	(0.009)	(0.007)	(0.013)	(0.007)
Asset Index	-0.004	0.026**	-0.011	0.026	0.002	0.023
	(0.011)	(0.011)	(0.012)	(0.016)	(0.017)	(0.015)
Being Orphan	0.057**	0.049*	0.039	0.078*	0.071**	0.030
	(0.024)	(0.028)	(0.026)	(0.043)	(0.036)	(0.036)
N^{o}. older siblings	-0.001	-0.001	-0.005	-0.004	0.004	0.001
	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)	(0.005)
N^{o}. of younger	0.007**	0.002	0.001	0.007	0.012**	-0.004
siblings	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors of marginal effects are calculated by the delta method and reported in parentheses. Marginal effects of regional dummies not shown. N^{o} observations: All: 2301; Men: 994; Women: 1307; Pseudo R^2 :All, 0.1497; Men, 0.1598; Women, 0.1594

Table A.1.2 (continued)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Community Characteristics						
Access-Primary School (1=yes)	-0.027	-0.150***	0.033	-0.159***	-0.092***	-0.146***
	(0.021) [a]	(0.035)	(0.022)	(0.051)	(0.034)	(0.046)
Access-Secondary School (1=yes)	0.013	-0.009	0.032	-0.086**	-0.020	0.071*
	(0.022)	(0.027)	(0.021)	(0.035)	(0.037)	(0.039)
Access-Hospital(1=yes)	-0.070***	0.044	-0.033*	0.090*	-0.092***	0.008
	(0.018)	(0.033)	(0.017)	(0.047)	(0.029)	(0.044)
Rural residence 10yrs (1=yes)	0.142***	-0.087***	0.114***	-0.119***	0.134***	-0.051
	(0.026)	(0.027)	(0.029)	(0.037)	(0.041)	(0.039)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors of marginal effects are calculated by the delta method and reported in parentheses. Marginal effects of regional dummies not shown. N° observations: All: 2301; Men: 994; Women: 1307; Pseudo R^2: All, 0.1497; Men, 0.1598; Women, 0.1594.

Table A.1.3 Multinomial logit models excluding temporary migrants (Marginal effects)

	Rural	Urban
Individual Characteristics		
Gender (1=male)	-0.022** (0.010)	-0.011 (0.013)
Age	0.041** (0.017)	-0.019 (0.020)
Age squared	-0.001** (0.000)	0.000 (0.000)
Wolof	0.034 (0.023)	-0.026 (0.023)
Poular	0.027 (0.021)	-0.008 (0.024)
Serere	0.048* (0.025)	-0.086*** (0.031)
Diola	0.013 (0.031)	0.030 (0.031)
Missing ethnicity	0.029 (0.021)	-0.053** (0.022)
Household Characteristics		
Father's highest education	-0.006 (0.004)	0.008** (0.003)
Mother's highest education	-0.001 (0.005)	0.001 (0.005)
Asset Index	-0.005 (0.008)	0.032*** (0.010)
Being Orphan	0.033* (0.017)	0.055** (0.024)
No. older siblings	-0.003 (0.003)	-0.002 (0.003)
No. of younger siblings	0.000 (0.002)	-0.003 (0.003)
Community Characteristics		
Access-Primary School (1=yes)	-0.016 (0.019)	-0.141*** (0.029)
Access- Secondary School(1=yes)	0.011 (0.017)	-0.006 (0.024)
Access- Hospital(1=yes)	-0.012 (0.015)	0.028 (0.028)
Rural residence 10yrs (1=yes)	0.044** (0.019)	-0.039 (0.025)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors of marginal effects are calculated by the delta method and reported in parentheses. Marginal effects of regional dummies not shown. N^o observations: All: 2301; Men: 994; Women: 1307; Pseudo R^2:All, 0.1497; Men, 0.1598; Women, 0.1594

Table A.1.4 Multinomial logit models with interaction between asset index and rural origin (Marginal effects)

	All		Men		Women	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
Household Characteristics						
Father's highest education	-0.017***	0.011***	-0.012*	0.005	-0.022**	0.017***
	(0.006)	(0.004)	(0.007)	(0.006)	(0.009)	(0.005)
Mother's highest education	-0.005	0.005	0.005	-0.006	-0.011	0.010
	(0.009)	(0.006)	(0.009)	(0.008)	(0.014)	(0.008)
Asset Index	0.030	0.033**	0.027	0.046**	0.032	0.019
	(0.019)	(0.014)	(0.025)	(0.020)	(0.027)	(0.019)
Asset Index*Rural residence	-0.059***	-0.027	-0.065**	-0.082**	-0.053	0.005
	(0.023)	(0.024)	(0.029)	(0.039)	(0.034)	(0.031)
Being Orphan	0.065***	0.036	0.054*	0.065	0.079**	0.020
	(0.025)	(0.031)	(0.032)	(0.049)	(0.037)	(0.039)
No. older siblings	-0.002	-0.002	-0.008*	-0.007	0.004	0.001
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.005)
No. of younger siblings	0.008**	0.001	0.002	0.005	0.013**	-0.003
	(0.003)	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. [a] Standard errors of marginal effects are calculated by the delta method and reported in parentheses. Marginal effects of individual and community characteristics as well as regional dummies are not shown. N^o observations: All: 2301; Men: 994; Women: 1307; Pseudo R^2: All, 0.1497; Men, 0.1598; Women, 0.1594.

CHAPTER 2

THE IMPACT OF EARLYCHILDBEARING ON SCHOOLING AND COGNTIVE SKILLS AMONG YOUNG WOMEN IN MADAGASCAR

Abstract

Female secondary school attendance has recently increased in Sub-Saharan Africa; however, the higher likelihood of attending school after puberty has put girls at risk of becoming pregnant while attending school. Using a panel survey designed to capture the transition from adolescence to early adulthood, we analyze whether teenage pregnancy contributes to lower school attainment and cognitive skills among young women in Madagascar. We address the endogeneity between fertility and education decisions by instrumenting early pregnancy with the young woman's access to condoms at the community level, and her exposure to condoms since she was 15 years old. We control for an extensive set of community social infrastructure characteristics to deal with the endogeneity of program placement. Our instrumental variable results show that having a child increases by 42% the likelihood of dropping out of school and decreases the chances of completing lower secondary school by 44%. This school-pregnancy related dropout is associated with a reduction of 1.1 standard deviations in the Math and French test scores. These results are consistent with hazard model estimations: delaying the first birth by a year increases the probability of current enrollment by 5% and the Math and French test scores by 0.2 standard deviations.

2.1 Introduction

Adolescent pregnancy can have detrimental economic and social consequences. In developing countries, early childbearing is associated, not only with health risks such as maternal death and low birth weight, but also with low school attainment and productivity, and consequently intergenerational transmission of poverty. There is, however, a paucity of empirical evidence that establishes a causal impact of early fertility on young women's human capital formation in developing countries. In Sub-Saharan Africa, pregnancy-related school dropout has increasingly gained prominence in part due to the recent expansion of female school enrollment in the region. The greater likelihood for girls to attend school after puberty has put them at risk of early pregnancy while they are attending school (Lloyd and Mensh, 2008). Girls face complex fertility and schooling decisions with the added constraints of low availability of information on safe sexual practices and limited access to reproductive health services in low-income countries (Chong et al 2013).

This paper investigates whether early childbearing has a causal effect on school attainment and cognitive skills, measured by Math and French test scores, in Madagascar. This country offers an appropriate context for our research question. Female progression to secondary school has rapidly increased from 45% to 69% between 1998 and 2010 in Madagascar (WDI, 2012); however, 32% of girls between 15 and 19 years old have a child or are pregnant for the first time; 48% of women age

18 are mothers or pregnant (DHS, 2009).[19] Indeed, Madagascar is among the top 10 developing countries with teenage pregnancy rates above 20% (UNPF, 2013).

Moreover, family planning prevalence in Madagascar is only 29% among women between 15 and 49 years old and there is no access to safe abortion[20].

The challenge of addressing the endogeneity between education and fertility is a result of the possibility that these two are joint decisions; for example, young girls may have strong preferences for education and labor market success and therefore less preference for children. Using a panel data survey designed to capture the transition from adolescence to adulthood, we use an instrumental variable approach to address this endogeneity. We estimate two sets of models: in the first, we instrument the young woman's early pregnancy with her "access to condoms", defined as the availability of condoms in the community where she lives; and in the second, we define the instrument in terms of "exposure to condoms", measured by the number of years during which the young woman has had community-level access to condoms since she was 15 years old. This age cut off is reasonable since the average age of first birth in our sample is 18[21]. Using the same identification strategy, we also estimate the age of first birth in the first stage using Weibull hazard models. This specification allows us to correct for the right censoring problem of women in our sample. The

[19] The total fertility rate is still at 4.8 children per women according to the 2008/09 Demographic Health Survey (DHS).

[20] In Madagascar abortion is illegal. However, some estimates of abortion rates are at 1 per 10 live births. Abortion complications are one of the major contributors to maternal death in the country (Sharp et al., 2011).

[21] The age of sexual initiation among Malagasy women is 17.4 (DHS,2009)

hazard also has the advantage of enabling us to estimate the impact on education of postponing the first pregnancy by one year.

The idea behind this identification strategy is that access (exposure) to condoms lowers fertility control costs among young women and affects their schooling decisions through the reduction of early pregnancy rather than through a direct effect. We show different robustness checks to support this argument including a placebo test in which access to condoms does not have a statistically significant direct effect on the school outcomes of young men. Additionally, we show that there is no evidence to support any non-random program placement of condoms. The reason we use condoms as an instrumental variable for early fertility, instead of other family planning methods such as pills or injectables, is that the latter are primarily used to space children within the family rather to postpone the first birth. There are also social norms that discourage girls of high school age from going to family planning clinics and seeking contraceptive access from the formal health care establishment. This is in contrast with condoms which are readily accessible among school girls; 40% of condom distribution is done through stores and 20% through pharmacies (DHS, 2009). Additionally, condoms are also a policy target employed to prevent Sexually Transmitted Infections (STIs) which prevalence is very high in Madagascar. In recent years, there has been an effort from the government and NGOs to increase condoms access and use among vulnerable populations (Glick et al, 2009).[22]

[22] The low HIV/AIDS prevalence in Madagascar despite a high prevalence of STIs (one of the highest in Sub-Saharan Africa), relative high rate of sexual partner change and early sexual initiation is considered an anomaly, but still a threat for public health (Sharp and Kruse, 2011).

One concern with this identification strategy is the possibility of non-random program placement: condom programs are potentially located in communities where teen pregnancies are the highest or where the population is more inclined to use contraception (Portner et al, 2012; Molyneux et al, 2000). To address this potential issue, we control in our models for an unusually complete set of social infrastructure community variables, both contemporaneous and lagged, which are available in the community surveys of the panel survey. Additionally, we estimate a linear probability model of the access to condoms at the time of survey on 2006 community-level fertility measures and a range of community covariates. We also estimate the same model but instead of fertility, we control for variables that capture the size of the population and poverty rates in 2001, a period over 10 years before the survey of the women used in our study. We do not find any evidence of non-random program placement in these models. Furthermore, we also control in our models for childhood socioeconomic characteristics when the girls were in the age range of 13 to 15 in order to isolate the effect of pregnancy from the effect of poverty conditions when women were young adolescents.

Our findings show that teenage pregnancy has a negative causal effect on schooling attainment and cognitive skills among young women in Madagascar. Our instrumental variable results show that early childbearing increases the likelihood of dropping out of school by 42%. Also, it decreases the chances of completing lower secondary school (i.e. completing more than 9 years of education) by 44%. These findings suggest that schooling and pregnancy are *non-compatible*, a result that is consistent with other empirical evidence in African contexts.[23] We also find that this early departure from school due to pregnancy has a detrimental impact on young women's cognitive skills: teenage pregnancy decreases by 1.1 the standardized scores of Math and French. Consistently, the results from the hazard models suggest that postponing first birth by one year has comparable gains in school attainment and cognitive skills.

Unlike in developing countries, the socioeconomic effects of teenage pregnancy in the United States (U.S.) have been extensively researched. A series of empirical strategies have been used to identify causal impacts and to deal with the systematic differences between mothers and non-mothers. These include employing sibling fixed effects which compares teen mothers to their childless sisters (Geronimus and Korenman, 1992), natural experiments that use miscarriages as an instrument of early fertility (Hotz, McElroy and Sanders, 2005; Ashcraft and Lang, 2006; Fletcher and Wolfe, 2009) as well as other instrumental variables for early fertility such as age of menarche, abortion and contraception rates (Ribar, 1994 and Keplinger et al. 1999) and propensity score matching methods within school attended to construct an appropriate counterfactual group for teenage mothers (Levine and Painter, 2003). There is no consensus over whether teenage pregnancy has a causal effect on poor school attainment, labor market outcomes and the probability of being a welfare assistance recipient. Except for Keplinger et al (1999), most of the studies in the U.S.

[23] See Duflo et al (2012) and Ozier (2011) for empirical evidence in Kenya and Baird et al (2011) for Malawi.

have found that the impact of adolescent pregnancy on these outcomes is smaller than the one implied by OLS regressions and sometimes is not statistically significant.[24]

In the context of developing countries, to the best of our knowledge, there are very few studies that rigorously establish a causal relationship between early pregnancy and school outcomes. Azevedo et al. (2012) use data on miscarriages as instrument for the timing of pregnancy in Mexico and find that a younger age of first birth does not have adverse effects either on education or on employment. In contrast, Arceo et al (2012), using a propensity score matching to construct a counterfactual group for the young mothers, find that teenage pregnancy decreases the years of schooling in Mexico. Employing the same methodology, Ranchod et al. (2011) find that high school completion in South Africa is driven more by socioeconomic conditions than by early pregnancy

Our paper, therefore, contributes to the limited empirical evidence of the impact of early childbearing on socioeconomic outcomes in developing countries, particularly in Sub-Saharan Africa. Our study is the first to show empirical evidence in a low-income and high-fertility country, in contrast to the aforementioned studies that have addressed this question in middle-income countries where, for example, social attitudes to teenage pregnancy, and institutions to help young women and their families cope with a teenage birth, may be more developed. Second, to our knowledge there is no empirical evidence of the effect of teenage pregnancy on cognitive skills. There are few cross-sectional studies in South Africa that have shown an association between test scores performance and fertility (Thomas, 1999) or the initiation of

[24] For a discussion of the recent empirical studies see Fletcher and Wolfe (2009)

sexual activity (Marletto, 2008), but they have not established a causal effect of early pregnancy on cognitive skills. Third, using access/exposure to condoms, as an instrumental variable (IV) rather than propensity score and miscarriages, allows us to infer the potential economic consequences of population policies that aim to decrease young women's costs to control fertility in low-income countries.

Indeed, our findings show that during the transition from adolescence to adulthood, reducing early childbearing or delaying the age of first birth generates substantial gains in education and cognitive skills among young women in Madagascar. Therefore, access to family planning and sexual reproductive health services for young women not only can prevent poor pregnancy outcomes but also can have a potential role in enhancing young women's education opportunities and increase their accumulation of human capital.

This paper is organized as follows. Section 2 describes the panel data set and the context of Madagascar. Section 3 presents the empirical methodology using access and exposure to condoms as instrumental variables, including the results of the hazard models to measure the impact of age of first birth on the school outcomes. Section 4 discusses the results, while the last section concludes and discusses policy implications.

2.2 Data Description and Context

The Madagascar Life Course Transition of Young Adults Survey 2011-12 re-interviewed a cohort of 1749 young adults, 859 of them women, who were 21-23 years old at the time of the survey. The survey gathered detailed socioeconomic information on them, their spouses, and the households in which they reside at the time of survey, as well as their communities. The survey also collected detailed retrospective event histories on the cohort members regarding schooling, fertility, employment, marriage, health as well as on the range of economic and life-course events, and experiences going back to 2004 in Madagascar, the year in which the cohort members were last surveyed in the *Enquête sur la Progression Scolaire à Madagascar* (EPSPAM) survey.[25] This 2012 data include tests of cognitive skills in Math and French collected for all the cohort members even if they were not attending school at the time of the survey. The 2011-12 also questioned community leaders, teachers and health personnel as to the availability of social and economic infrastructure and services at the community level, including information on family planning services, as well as the date since these services were available in the community. We complement the information at the community level with the 2001 and 2007 commune censuses that feature a wide range of information about all the villages in Madagascar, including information on the basic public services and infrastructure.

For the 2012 sample of 859 women cohort members between 21 to 23 year old, we have detailed fertility and education history information, including test scores from 2012. Some descriptive statistics on these women's education are summarized in Table 1. We observe that 54% (466) of the women in the sample have given birth to at least one child; we call this group of women *"Ever mothers"*. We call *"Non-Mothers"* their female counterparts in the cohort who have not yet given birth. In our sample, the average age of first birth is 18 years (standard deviation 2.12).

[25] Around 90% of the 2004 cohort members were followed in 2012. This attrition rate (around 10%) is very small compared to other panels in Sub- Sahara Africa.

Table 2.1 Education and Cognitive Performance for Mothers and Non-Mothers

	Non Mothers	Ever Mothers	All
2012 Education variables			
% School Enrollment	34.00	3.27	17.39
Years of Education	9.25	6.20	7.60
	(3.74)	(3.18)	(3.77)
% Some Primary school	13.74	29.18	22.12
% Completed Primary	10.69	18.67	15.02
% Some Lower Secondary	9.67	19.96	15.25
% Completed Lower Secondary	14.25	14.81	14.55
% Some Upper Secondary	10.18	7.51	8.73
% Completed Upper Seconday	23.41	4.94	13.39
% Some University	17.56	1.93	9.08
Cognitive skills			
2012 Math Test Score	16.43	11.78	13.97
	(8.12)	(7.10)	(7.94)
2012 French Test Score	12.28	7.92	9.98
	(6.22)	(5.75)	(6.35)
No of Observations	393	466	859

Notes: Standard deviations in parentheses. Girls without any education represent 1.86 % of the sample. This percentage is not shown in the table 2.1.
Differences among groups statistically significant at 1%

Table 2.1 shows substantial differences in schooling and cognitive performance between these two groups. While 34% of the *non-mothers* still attend school, only 3% of the *ever-mothers* are enrolled. We calculate the difference between the age of awareness of conception and age of dropping out of school, and classify the young women according to the timing of these two decisions of fertility and education (See figure 2.1). We find that almost 24% of the sample, or 46% of the young mothers got pregnant while they were in school.[26] In contrast, 30% of the girls drop out of school but they do not get pregnant at least until the time of the survey. Also, it is noteworthy that 27% of the young women drop out much sooner than their first birth, indicating that there is no overlapping between their fertility and education decisions. Finally, we also observe that only 16% of the girls are still attending school at the time of the survey and are "non-mothers". A very negligible proportion of the sample (2%) are "ever mothers" and currently enrolled in school which suggests the difficulty of continuing education once the young women have their first child.[27]

These patterns are consistent with the years of education completed among the two groups as shown in Table 2.1. While the group of *ever-mothers* completed 6.2 years of schooling, the corresponding figure for *non–mothers* is 9.25. This difference is reflected in the data on the progression through school. Among the group of *ever-mothers*, only 5% completed upper secondary while this percentage is almost 5 times larger among the *non-mothers*. Also, 17% of the women who have not yet had their first birth have some university education while this percentage among young mothers is negligible.

---

[26] Given that we have an exact women's children date of birth but not a calendar of their pregnancy, we calculate the age of awareness of conception as the age of the first birth minus eight months of pregnancy.

[27] In field work visits, different stakeholders in the education sector pointed out the fact that school girls who get pregnant are socially and sometimes pressured by the school principal, to leave the school to avoid a reputational cost for the school. Also, in Madagascar there is no law or regulation in the education sector to address the problem of pregnancy school-dropout, although the issue is recognized by the authorities.

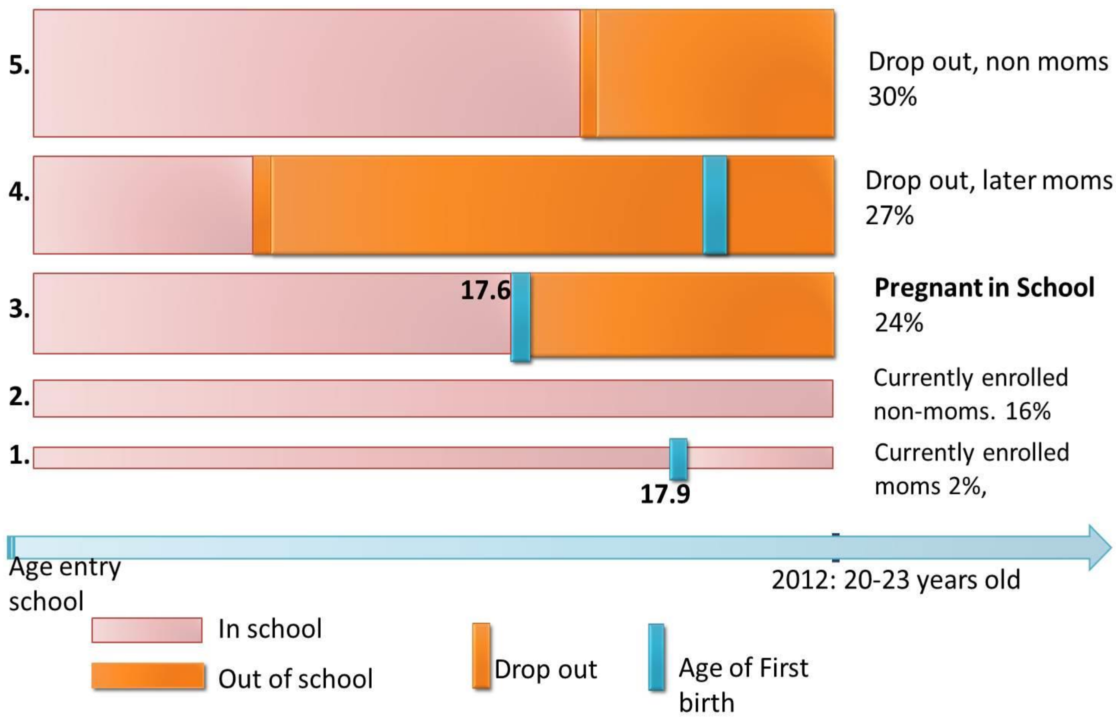
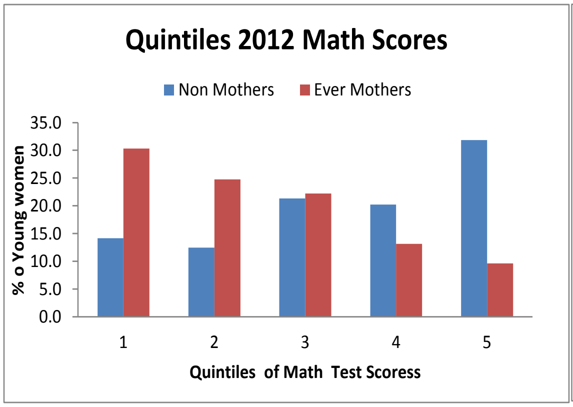


Figure 2.1 School enrollment, dropping out and pregnancy

Additionally, we report in Table 2.1 that the *non-mothers* have on average better performance in the 2012 Math and French test scores, compared to the *ever-mothers*. This is a reflection, in part, of the fact that the former group stays in school longer. The share of young women in the upper quintiles of the Math and French test scores distribution is far greater for those who are not yet mothers in 2012 than for those women who have given birth by 2012 (See Figure 2.2).



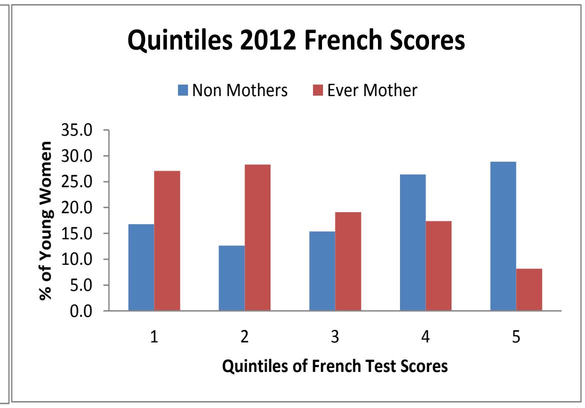


Figure 2.2 Quintiles of Math and French test scores by Ever Mothers and Non Mothers

Regarding the use of family planning among the young women in our cohort, the data shows that 31% of them use at least one method of contraception (modern and/or traditional). As Table 2.2 shows, there is a larger group of family planning users among the "*ever mothers*", compared to the "*non-mothers". This is consistent with the fact that almost 40% of women in Madagascar use family planning for the first time when they have already at least one child (DHS, 2009). Among the family planning users, 37 % have primary school, 38% have lower secondary and the rest have upper secondary or higher education. There is no evidence of a positive correlation between the young women's level education and their use of family planning in our sample. In terms of access to family planning services, defined as the existence of these services in the community where the young woman lives, Table 2.2 shows that the group of *non-mothers* have more access to family planning services, specifically, to pills and condoms than the group of *ever-mothers*.

Table 2.2 Access and Family Planning Use Among Young Women

	Non Mothers	Ever Mothers	All
% Family Planning Use	18.07	42.27	31.2
% 2012 FP services access	91.09	80.9	85.56
% 2012 Pills access	83.21	73.61	78
% 2012 Condom access	84.48	69.1	76.1
N	393	466	859

Note: The differences between non-mother and ever mothers on the FP variables of the table are statistically significant at 1% level.

In the next section we describe the identification strategy we employ to study the impact of early childbearing on school attainment and cognitive skills among young women in Madagascar.

2.3 Empirical Strategy

Fertility and education decisions might be endogenous; i.e omitted variables such as ability, and motivation can be correlated with these two decisions. Therefore, OLS estimations of the early childbearing impact on schooling and cognitive skills might be biased. The ideal experiment to analyze this impact would be to randomly assign pregnancies among girls and compare their educational outcomes after girls have given birth. This is an impractical and unethical research design. To address this endogeneity issue and exploiting the available information in the survey, we use the young woman's access and exposure to condoms at the community level as an Instrumental Variables (IV) for her early childbearing. Access is defined as the availability of condoms in the community where the young woman lives.

We use condoms, rather than another family planning method, as an IV for young woman's first birth for the following reasons. First, condoms are considered a key policy target employed to prevent Sexual Transmitted Infections (STIs) and pregnancy among young women (Chong et al 2013). Second, while injectables, pills, and condoms are widely known to women in Madagascar,[28] the first two are more common among women who already have children; thus, pills and injectables are primarily used to space children within the family rather than to postpone the first birth. In fact, 38% of women use family planning for the first time when they have already at least one child (DHS, 2009). [29] Consistently, condoms are more used by single than married women in Madagascar (Glick et al, 2009). Third, in contrast to pills and injectables, condoms are not perceived to have negative secondary health effects.[30] Using condoms as an instrumental variable avoids the problem of accounting for social norms and misconceptions about the use of contraception. Fourth, as expressed by NGO workers and government agents during our field work visits (2012), school girls have the stigma of going to family planning centers to get the injections or pills, whereas condoms are more easily accessible in this target population; 40% of them are available in stores and 20% in pharmacies (DHS, 2009).[31]

[28] According to the 2009 DHS, 87.9% of women have heard of pills, 89% of injectables and 85% of condoms.

[29] Moreover, 11% of women used family planning for the first time when they have already 4 children or more (DHS, 2009).

[30] According to 2009 DHS, 18% of the women, aged 15-49, who are in union do not use family planning due to the potential secondary effects, while in the case of condoms, no negative health effects are expected.

[31] We also tried our models with access to pills since this information was also collected in our data. The F-stat of the first stage was only 2 consistent with the reasons why condoms are better instrument for first pregnancy.

The IV approach involves estimating a two-stage model of the following form:

$$Y_i = a + \beta' EverMother_i + \pi' Age_i + \rho' X_i + \theta' C_i + u_i \quad (1)$$
$$EverMother_i = \mu + \tau' Age_i + \delta' Zi + \gamma' X_i + \varphi' C_i + v_i \quad (2)$$

Where $\boldsymbol{Y_i}$ represents distinct educational outcomes in 2012 : i) current enrollment, ii) years of schooling, iii) a dummy variable for completing secondary school (i.e., having 9 or more years of education), and iv) standardized French and Math test scores.[32] Age_i is a set of dummies for women's age cohort, X_i is a vector of women's parents' socioeconomic variables and C_i is a vector of extensive community variables to control for potential placement endogeneity of condoms, as we explain in more detail later. We restrict the estimation to those girls who quit school at an age greater than 13 years or older, thus excluding 10% of the total 859 women in the sample[33]. We do so to guarantee that girls are attending school at the minimum age at which they might be at the risk of pregnancy. Table 2.3 includes a summary statistics of the variables used in the estimation.

[32] The standardized scores are constructed by subtracting the mean and dividing by the standard deviation of the sample.

[33] We choose the age of 13 years because, according to Walker et al. (2011), this is the median age that girls drop out of school in Madagascar in 2004.

Table 2.3 Summary Statistics of the Variables in the IV Models

	N	Mean	Std Dev
Dependent Variables			
Ever_mother_d (Y=1)	778	0.526	0.500
Current Enrolment (Y=1)	778	0.189	0.392
Years of Schooling	778	8.076	3.577
Z Scores of french	703	0.076	0.985
Z Scores of Math	712	0.074	0.981
Parents' variables			
Asset Index 2004	758	0.106	0.997
Mother is alive (Y=1)	778	0.905	0.294
Father alive(y=1)	778	0.823	0.382
Mother 's years of educaiton	774	4.903	3.578
Father's year of educaiton	774	5.598	3.944
Community variables at the time of the survey			
Community health Center (CSB2)	778	0.636	0.481
Community Hospital (CHD1)	778	0.135	0.342
Upper Secondary (y=1)	778	0.614	0.487
Piped Water (Y=1)	778	0.554	0.497
Access to weekly market (Y=1)	778	0.614	0.487
Access to paved road all year(Y=1)	778	0.422	0.494
Urban Indicator	778	0.289	0.454
Community variables at 10 years old			
Electricity at 10 yrs old	778	0.490	0.500
Upper Secondary at 10 yrs old	778	0.413	0.493
CSB2 at 10 yrs old	778	0.541	0.499
Electricity at 10 yrs old	778	0.276	0.447
Remotness index 2001	778	2.377	1.345

Notes: Not shown dummies for cohort age and regions

We estimate two sets of models. In the first, the instrumental variable $\boldsymbol{Z_i}$, is a dummy variable for whether the young woman has “access to condoms” in the community where she lives. It is worth noting that access is not a self-reported measure since this question in the survey was answered by community leaders. Also, as mentioned earlier, condoms are not distributed in schools. Therefore, we are not concerned that young women are receiving condoms while attending school. In addition, condoms are free or their price is heavily subsidized by the government or NGOs, therefore, price is not a factor defining access to this family planning method.[34] Given that we have information on the specific year that condoms were available in a particular community, in our second set of models, we use the "exposure to condoms" as an IV. Exposure is defined as the number of years that a girl has had access to condoms at the community level since she was 15 years old. In our sample of communities, the median year that condom distribution started is 2000 and the average years of exposure since girls were 15 years old is 4.8. This age cut-off seems reasonable, not only because reproductive health programs of NGOs and government focus on young people starting at 15 years and older, but also because the median age of sexual initiation among Malagasy women is 17.4 (DHS, 2009) and the average age of first birth in our sample is 18.[35] We estimate 2-SLS models and IV-probits for binary dependent variables outcomes when instrumenting with the exposure to condoms. [36]

In both equations, we control for young women cohort's age dummies, $A_{i.}$ We also include $\boldsymbol{X_i}$, a set of young women's parents control variables: a dummy for whether the parents were alive or not by the time of survey (2012)[37], parents'

[34] According to the USAID sources, the price of a box containing 8 condoms is 400 ariary (0.15 USD)
[35] As a robustness check, we also estimate our IV models with a different measure of exposure; i.e since age 10 and 12. Although these instruments are negatively correlated with the likelihood of pregnancy, the F-statistics of the first stage are lower than the one of the exposure since age 15. Models are upon request.

[36] Lineal IV Models (2-SLS)using exposure to condoms are available upon request. These indicate similar results than IV probit models

[37] We also try some specification with the time varying variable of the parents alive at age of and 15 but they did not have a different statistically effect than the 2012 parent alive variable in the models

education, and an asset index constructed from the earlier round of the survey in 2004, when the women were on average 15 years of age.[38] These variables are important since other studies in the region (see for example, Ranchod et al. 2011) indicate that girls' educational attainment is driven more by socioeconomic conditions than by early pregnancy. Also, the inclusion of the asset index and parents education is relevant given their importance as determinants of cognitive skills and school dropout in Madagascar (Glick et al., 2009 and Walker et al., 2011).

One concern with the IV strategy outlined above is that access to condoms might be related to the level of social infrastructure and public services provision at the community level. This potential for a non-random placement of the condoms has been discussed by Portner et al. (2012) and Molyneux et al. (2000). Condoms might be potentially located in communities where teen pregnancies are highest or, conversely, where the population is more inclined to use contraception. This might have large indirect effects on young women's joint fertility and education decisions (Angeles et al 2005; Glick et al. 2011; Portner et al. 2012). To address this issue, we include in the equations 1 and 2 above the variable *Ci*, an extensive set of social infrastructure variables that allows us to identify the effect of access to condoms (exposure to condoms) conditional to these variables. We use variables from the 2012 community survey and the 2007 commune census to control for access (defined as the availability in the community) to upper secondary school, district hospital health center (CHD1), community health center (CSB2), electricity, piped water, weekly market and paved

[38] The asset index was created by using ownership of durable goods such as radio, TV, refrigerator and bicycle motorcycle or car as well as the source of drinking water and toilet facilities of the dwelling. For details on the asset construction see Glick et al. (2012) and Sahn and Stifel (2003)

roads. Furthermore, we control for time varying community covariates by including access to secondary school, electricity and community health center (CSB2) when the girl was 10 years old. From the 2001 community census, we include a remoteness index which is created using factor analysis and information on community distances to the main social infrastructure services and transportation.[39] In combination, these control variables allow us to uncover some of the unobservable characteristics at the community level that might be related to the access to condoms. The inclusion of these community variables avoids implementing IV fixed effects models at the community level.

In addition to controlling for C_i in our models, we check for potential non-random placement of condoms by estimating a linear probability model where we predict the probability that a community has access to condoms at the time of survey not only in function of the community variables described earlier, but also in function of two different 2006 fertility variables at the community level: number of births and number of women who died during or immediately after child delivery. This last variable can be a good proxy of adolescent pregnancy since maternal mortality is higher among young women (UNPF, 2013). This fertility information comes from the 2007 commune census and it is not available for all the 73 communities of our sample. Table 2.4 shows that the 2006 fertility variables are not statistically significant and their coefficient are very small suggesting that condoms are not placed where teen

[39] The remote index includes health services, banks, post offices, schools, taxis, courts, markets, inputs, extension services, and veterinarians as well as access to national and provincial roads, utilities, media and other markets and several measures of access to transport.

and/or total fertility is higher. None of the other community covariates are statistically significant; therefore, we fail to reject the hypothesis of a random placement.

Furthermore, we estimate the same linear probability model of access to condoms but instead of 2006 fertility, we control for population and poverty variables from the 2001 community census, around 10 years before the 2012 survey. We are able to do this exercise for 71 of our 73 communities in 2012 survey. Table 2.4 shows that none of these covariates are statistically significant; indicating that there is no evidence for any non-random placement.[40] We also tried to instrument early pregnancy separately with the remoteness index and the access to CSB2 to test if the access to condoms effect on fertility was driven by other characteristics of the communities, but we did not find any statistical evidence to support that hypothesis (models not shown).

[40] Similarly, we estimate linear probability models of access to family planning services on the 2006 fertility and 2001 population and poverty controls as well as the other community variables. We do not find evidence of a non-random program placement. Models are available upon request.

Table 2.4 Linear Probability Models for Access to Condoms

	Access to Condoms (1)	Access to Condoms (2)	Access to Condoms (3)
Urban (Yes =1)	0.226	0.170	0.0828
	[0.242]	[0.241]	[0.256]
Electricity (Y=1)	0.137	0.133	0.126
	[0.0989]	[0.107]	[0.110]
Piped Water (yes=1)	0.105	0.106	0.174*
	[0.0956]	[0.0970]	[0.104]
Upper Secondary (Yes =1)	0.0242	0.0116	0.0539
	[0.151]	[0.149]	[0.148]
Community Health Center (CSB2)	0.321	0.318	0.300
	[0.207]	[0.204]	[0.205]
Hospital -CHD1 (Y=1)	0.134	0.119	0.165
	[0.120]	[0.114]	[0.132]
2001 Remotness Index	-0.0243	-0.0259	-0.0135
	[0.0520]	[0.0565]	[0.0550]
Access to weekly market	0.158	0.124	0.0717
	[0.137]	[0.131]	[0.140]
Access to Paved Road	0.119	0.151	0.122
	[0.114]	[0.118]	[0.118]
No Cyclones 2002-11	-0.0455	-0.0329	-0.0216
	[0.0571]	[0.0579]	[0.0566]
2006 No. of Births	-0.0000277		
	[0.0000310]		
2006 No. of Women dead after/immediately after child delivery		0.00144	
		[0.00921]	
2001 Log Population			0.0157
			[0.0726]
2001 % Poor people			-0.00190
			[0.00236]
N	68	66	71
adj. R-sq	0.158	0.154	0.125

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively.
Robust Standard errors reported in parentheses. Regional dummies not shown

A final point that must be addressed is that identification of the IV model requires a strong correlation between "access to condoms" (or "exposure to condoms") and the endogenous variable "ever mother". Table 2.5 shows the results of the first stage regression using access to condoms with and without the set of control variables described earlier. Table 2.6 presents the same results, but this time using the instrument of "exposure to condoms". We observe in Table 2.5 that having access to condoms at the community level, without controlling for any of the covariates, decreases by 26% the probability of being a mother, at the 1% statistical significance level (F-stat=38.75). Once we include the complete set of household and community control variables (column 5) this effect decreases to 18%, but it is still statistically significant at the same level (p value =0.001) with an F-stat of 11.36, above the Staiger and Stock criteria for weak instruments.

Table 2.5 First Stage using Access to Condoms as an IV

	(1)	(2)	(3)	(4)	(5)
Condom Access (Y=1)	**-0.262*****	**-0.193*****	**-0.199*****	**-0.182*****	**-0.179*****
	[0.0395]	**[0.0441]**	**[0.0497]**	**[0.0514]**	**[0.0530]**
2004 Asset Index		-0.0529***	-0.0440*	-0.0171	-0.0122
		[0.0203]	[0.0241]	[0.0261]	[0.0273]
Mother is alive (Y=1)		-0.0496	-0.0386	-0.0417	-0.0390
		[0.0628]	[0.0616]	[0.0610]	[0.0610]
Father is alive (Y=1)		-0.00782	-0.0187	-0.0255	-0.0188
		[0.0452]	[0.0449]	[0.0449]	[0.0453]
Mother's years of education		-0.0141**	-0.0147**	-0.0155**	-0.0147**
		[0.00601]	[0.00602]	[0.00600]	[0.00605]
Father's years of education		0.00000143	0.00152	0.00107	0.000587
		[0.00572]	[0.00571]	[0.00563]	[0.00567]
Community health Center (CSB2)			0.0579	0.0494	0.0808
			[0.0482]	[0.0744]	[0.0888]
Community Hospital (CHD1)			-0.0353	0.00540	-0.000210
			[0.0528]	[0.0539]	[0.0577]
Upper Secondary (y=1)			-0.105**	-0.0890	-0.0965
			[0.0449]	[0.0594]	[0.0642]
Piped Water (Y=1)			0.0751*	0.0832**	0.0951**
			[0.0398]	[0.0417]	[0.0438]
Access to weekly market (Y=1)			-0.0666	-0.0903*	-0.104**
			[0.0443]	[0.0466]	[0.0481]
Access to paved road all year(Y=1)			0.0551	0.0581	0.0683
			[0.0411]	[0.0443]	[0.0460]
Access to Electricity (Y=1)			0.00814	0.0738	0.0597
			[0.0475]	[0.0524]	[0.0585]
Upper Secondary at 10 yrs old				0.0288	0.0486
				[0.0630]	[0.0705]
CSB2 at 10 yrs old				-0.0288	-0.0276
				[0.0648]	[0.0660]
Electricity at 10 yrs old				-0.208***	-0.222***
				[0.0647]	[0.0761]
Remoteness index 2001				0.00177	-0.00527
				[0.0201]	[0.0207]
Urban					0.0135
					[0.108]
Regional dummies					Y
N	778	750	750	750	750
Fstat	**38.7516**	**19.2553**	**16.0394**	**12.4651**	**11.3694**
R-sq	0.0719	0.0969	0.1181	0.1312	0.136

Notes: Robust standard errors in brackets * p<0.10, ** p<0.05 *** p<0.01

All the models (1-5) include cohort age dummies. Model 5 inlcudes the Regional dummy variables not shown

Similarly, Table 2.6 indicates that having one extra year of exposure to condoms at the community level since the age of 15 decreases the probability of having children by 3.7% when no covariates are included in the estimation. This result is statistically significant at the 1% level (F-stat=44.2.). Once we control for the full set of variables, this estimate decreases to 2.3%, but remains statistically significant at the 1% level (p-value =0.001) with an F-stat value of 11.84.These F-statistics do not indicate a problem of weak instruments. This measure of exposure is consistent with the results of access to condoms, since 2.3% multiplied by the median exposure (4.8) is approximately equal to the point estimate of the access to condoms (18%). It should be noted that the effect of “access to condoms” and “exposure to condoms” on the endogenous variables does not change significantly in magnitude once we control by the 2012 and time-varying community variables and regional dummy variables. This robustness of the magnitude supports the hypothesis that there is no strong relationship between the access to condoms and the social infrastructure at the community level. For the rest of the results, we keep the complete set of control variables found in column 5 of both tables 2.5 and 2.6.

Table 2.6 First Stage using Exposure to Condoms as an IV

	(1)	(2)	(3)	(4)	(5)
Condom Exposure 15 yrs	**-0.0368*** **	**-0.0277*** **	**-0.0257*** **	**-0.0230*** **	**-0.0234*** **
	[0.00553]	**[0.00606]**	**[0.00640]**	**[0.00659]**	**[0.00680]**
2004 Asset Index		-0.0515**	-0.0448*	-0.0186	-0.0135
		[0.0203]	[0.0239]	[0.0258]	[0.0270]
Mother is alive (Y=1)		-0.0533	-0.0412	-0.0432	-0.0408
		[0.0632]	[0.0621]	[0.0614]	[0.0614]
Father is alive (Y=1)		-0.0119	-0.0229	-0.0291	-0.0219
		[0.0454]	[0.0451]	[0.0450]	[0.0453]
Mother's years of education		-0.0146**	-0.0149**	-0.0156***	-0.0147**
		[0.00603]	[0.00606]	[0.00605]	[0.00610]
Father's years of education		-0.000134	0.00124	0.000876	0.000345
		[0.00574]	[0.00574]	[0.00567]	[0.00571]
Community health Center (CSB2)			0.0386	0.0343	0.0745
			[0.0468]	[0.0742]	[0.0888]
Community Hospital (CHD1)			-0.0373	0.00325	-0.00206
			[0.0528]	[0.0540]	[0.0577]
Upper Secondary (y=1)			-0.0945**	-0.0828	-0.0903
			[0.0450]	[0.0599]	[0.0649]
Piped Water (Y=1)			0.0647	0.0715*	0.0854**
			[0.0393]	[0.0408]	[0.0428]
Access to weekly market (Y=1)			-0.0759*	-0.0959**	-0.112**
			[0.0441]	[0.0465]	[0.0477]
Access to paved road all year(Y=1)			0.0444	0.0501	0.0607
			[0.0409]	[0.0443]	[0.0460]
Access to Electricity (Y=1)			-0.000436	0.0658	0.0509
			[0.0471]	[0.0523]	[0.0583]
Upper Secondary at 10 yrs old				0.0381	0.0559
				[0.0629]	[0.0700]
CSB2 at 10 yrs old				-0.0301	-0.0295
				[0.0651]	[0.0662]
Electricity at 10 yrs old				-0.204***	-0.223***
				[0.0641]	[0.0743]
Remoteness index 2001				0.00537	-0.00277
				[0.0199]	[0.0205]
Urban (Y=1)					0.0251
					[0.106]
Regional Dummies					Y
N	778	750	750	750	750
First Stage	**44.2365**	**20.8813**	**16.143**	**12.2032**	**11.8428**
R-sq	0.0719	0.0988	0.1178	0.1307	0.1365

Notes: Standard errors in brackets * p<0.10, ** p<0.05 *** p<0.01"

All the models (1-5) include cohort age dummies. Model 5 includes the Regional dummy variables not shown

In the results presented in table 2.5 and 2.6 above, we do not cluster the standard errors at the community level. We have reasons in favor and against to do so, given the size of our sample. Although the endogenous variable “ever mother” varies at the individual level, access to condoms varies at the community level. Therefore, fertility decisions of women living in the same community might be correlated. If this is the case, clustering will correct the implied underestimation of the standard errors (Bertrand, et al 2004). On the other hand, our sample includes 73 communities (above the critical level of 50 to cluster) but in each community we have less than 5% of the total sample and the distribution of individuals across communities is unbalanced. Rogers (1993) shows that in these cases clustering can do more harm than good[41]. We show in table A.2.1 of the appendix the first stage of the models using clustered standard errors controlling for all the individual, household and community variables. The F-stat decreases to 8.16 and 7.11 for access and exposure to condoms, respectively[42]. These F-stat magnitudes are under the rule of thumb for weak instruments; however, because the models are just-identified, the weak instrument bias towards OLS is not present (Angrist and Pischke, 2009). Also, Table A.2.2 shows the results of the second stage of the IV models using clustered standard errors. The statistically significance is only compromised in the case of standardized scores of Math; however, our main results are robust to this more conservative scenario.

[41] Indeed, we have 3 clusters with 2 individuals each. We estimate our models without these clusters and the results are robust to this exclusion.

[42] In these first stages, the point estimates of ‘access to condoms’ and ‘exposure to condoms’ are statistically significant at the 1% level , respectively ,in each regression the p-values are 0.006 and 0.008.

Furthermore, our results are robust to the estimation of the two stage models using IV-GMM and IV LIML suggesting that there is not weak instrument problem.[43]

2.3.1 Hazards Models Predicting the Age of First Birth

Using the same identification strategy, we estimate in the first stage a Weibull hazard model in which failure occurs when the young woman has her first child. This hazard model addresses the issue of right censoring since almost half of the young women in our sample have not yet had their first birth: for these women we only know that age at first birth is at least as high as the current age. Thus, estimating the age at first birth can be done by modeling duration (years) until the first birth:

$$h_j(t) = h_o(t)\exp\{\delta' Age_j + \beta' Z_j + \alpha' X_j + \rho' C_j\}$$

Where the hazard rate $h(t)$ is the probability of having the first birth at time (or age) t conditional on not having a child until t, *Age* is the birth cohort dummies, Z is the exposure to condoms, X_j and C_j are respectively the household and community characteristics described earlier. The term h_o is the baseline hazard that in a Weibull distribution is defined by: $ho(t) = pt^{p-1}$. [44] We choose this parameterization because in our sample we expect that the probability of having the first child increases with age (p> 1).[45] The Weibull model allows us to calculate an expected predicted

[43] The results of IV-GMM and LIML are upon request.

[44] We reject the test of proportionality for the Cox hazard which suggests a parametric hazard model.

[45] Examples of using the Weibull distribution to model the age of sexual initiation in Africa include Glick and Sahn (2008).

survival time; that is, an expected "predicted age of first birth-*PredAFB*"[46] which we use in the second stage to predict the school outcomes:

$$Y_i = a + \beta PredAFB_i + \pi' Age_i + \rho' X_i + \theta' C_i + u_i$$

Where *Yi* corresponds to the different school outcomes previously analyzed. Table 2.7 shows the hazard ratios for the main covariates using exposure to condoms.[47] We obtain the same qualitative results as when modeling the probability of "ever mother". Figure 2.3 shows the predicted hazard function after estimating the Weibull model that controls for access to condoms and the rest of covariates described earlier.[48] We observe that young women in communities where there is access to condoms have a lower risk of being pregnant, which confirms the validity of our identification strategy. Following our earlier discussion on the sensitivity of results to clustering, we present in Table 2.7 the results without clustering; however, assumptions on the standard errors do not affect the predicted age of first birth. Therefore, clustering or not will only affect the significance level of the parameters of interest in the second stage.[49]

---

[46] The expected mean age of first birth; i.e., the expected value of the survival time is given by :

$$\text{PredAFB} = \int_0^{\infty} S(t|X_j) \quad \text{where } S() \text{ is the survival function of the Weibull Distribution.}$$

[47] Results using access to condoms are qualitatively similar. We keep the specification with exposure to condoms since this instrumental variable is more appropriate for a duration model.

[48] Similar pattern is observed when using Kaplan Meier estimates by access to condoms.

[49] We have single spell continuous hazard models; one spell corresponds to one woman. Thus, the clustering issue has the same nature as the IV lineal models.

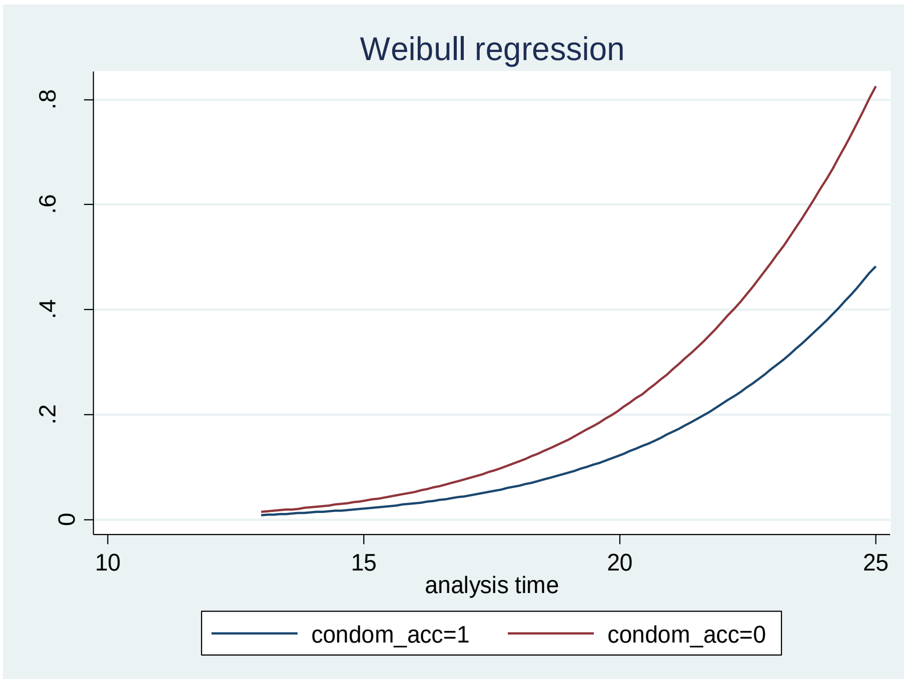
Table 2.7 First Stage for Age of First Birth

	Hazard ratio	Robust Standard error	z	P>|z|
Condom Exposure 15 yrs	0.94	0.019	-3.11	0.002
2004 Asset Index	0.92	0.083	-0.88	0.376
Mother is alive (Y=1)	0.87	0.164	-0.73	0.464
Father is alive (Y=1)	0.91	0.130	-0.65	0.513
Mother's years of education	0.95	0.019	-2.63	0.009
Father's years of education	1.00	0.018	-0.2	0.842
Community health Center (CSB2)	1.36	0.392	1.07	0.286
Community Hospital (CHD1)	0.94	0.178	-0.32	0.752
Upper Secondary (y=1)	0.75	0.157	-1.38	0.168
Piped Water (Y=1)	1.30	0.182	1.87	0.062
Access to weekly market (Y=1)	0.69	0.106	-2.4	0.016
Access to paved road all year(Y=1)	1.16	0.165	1.05	0.295
Electricity(Y=1)	1.16	0.206	0.82	0.41
Upper Secondary at 10 yrs old	1.14	0.270	0.54	0.587
CSB2 at 10 yrs old	0.87	0.181	-0.66	0.506
Electricity at 10 yrs old	0.46	0.117	-3.05	0.002
Remoteness index 2001	0.95	0.056	-0.95	0.341
Urban	1.19	0.396	0.53	0.593

p =7.14 (std error 0.22)
No of Observations = 750 ; Wald Chi2(27)= 113.62

Notes: Age Cohort and regional dummies not shown
Hazard ratios less than 1 decrease the risk of failure (ever mother) and opposite happens

As a robustness check, we estimate the Weibull hazard models changing the duration time from the age of 12 to the time of first birth or to the age in 2012 for the right-censored observations. The expected predicted age of first birth does not change significantly under this specification and neither do the second stage results. (See table A.2.5 in the appendix).[50]



Note: This is the predicted hazard function by access to condoms after estimating the Weibull model which controlls by the individual, household and community covariates

Figure 2.3 Weibull Model Estimations for Age of the First Birth

2.4 Results and Discussion

Table 2.8 reports the OLS and IV estimates of the early childbearing effect on: i) current enrolment, ii) years of education and iii) completion of lower secondary school, using a dummy variable for whether a young girl has completed 9 or more years of education. These outcomes are measured in the last wave of the survey (2012)

[50] Furthermore, we also estimate the hazard models with gamma distribution instead of Weibull and the results of the second stage model are qualitatively similar.

and among the girls who drop out from school at an age greater than 13. We present the OLS and IV models using both instrument variables, "access to condoms' and "exposure to condoms since age 15". We present in table 8 the average marginal effects for the binary outcomes, estimated from the IV probit models.

The OLS results indicate that having a child decreases by 27% the probability of being currently enrolled in school at the time of the survey. This estimate increases to 42% in both specifications of the IV model at the 5% significance level using "access to condoms" and at the 1% level in the IV- probit model using "exposure to condoms".[51] Although the estimates are larger in the IV specification than in the OLS model, this difference is not statistically significant.[52] Compared to the sample mean, this marginal decrease translates into a drop from 19% to 11% in the current enrollment. These findings suggest that there is a high opportunity cost in terms of forgone schooling for the girls who get pregnant. They also suggest that schooling and pregnancy are non-compatible or mutually exclusive as it has been shown in other African contexts, such as Kenya (Duflo et al 2012, Ozier 2011) and Malawi (Baird et al, 2011). In Madagascar, pregnant girls are commonly expelled from school *de facto* but there is no regulation justifying this practice. On the other hand, these results are different from those reported by Ranchod et al (2011) in South Africa. Although the

[51] To validate our results on school dropout , we construct a woman-year panel data using the age of dropping out of school and the age of first birth in order to analyze the effect of *"ever mother"* on current enrollment using a woman fixed effects specification. This estimation allows controlling for all the young woman's unobserved time invariant characteristics that might affect education and fertility decisions simultaneously. In this model (results not shown), early childbearing increases by 22% the likelihood of dropping out of school which is in line with our OLS and IV estimates.

[52] We are not able to reject the null hypothesis of exogeneity under the Haussmann and Durbin Watson tests.

authors find that teenage pregnancy statistically increases the school dropout[53] by 16% at age 20 or 22, they find smaller or negligible effects on high school graduation. Teenage pregnancy decreases high school graduation only by 5.9% by age 20 and 2.7% by age 22 and this latter effect is no longer statistically significant. This suggests that teen mothers can "catch up" in education, reflecting the possibility of policies facilitating their return to school

Table 2.8 Impact of Early Childbearing on School Attainment

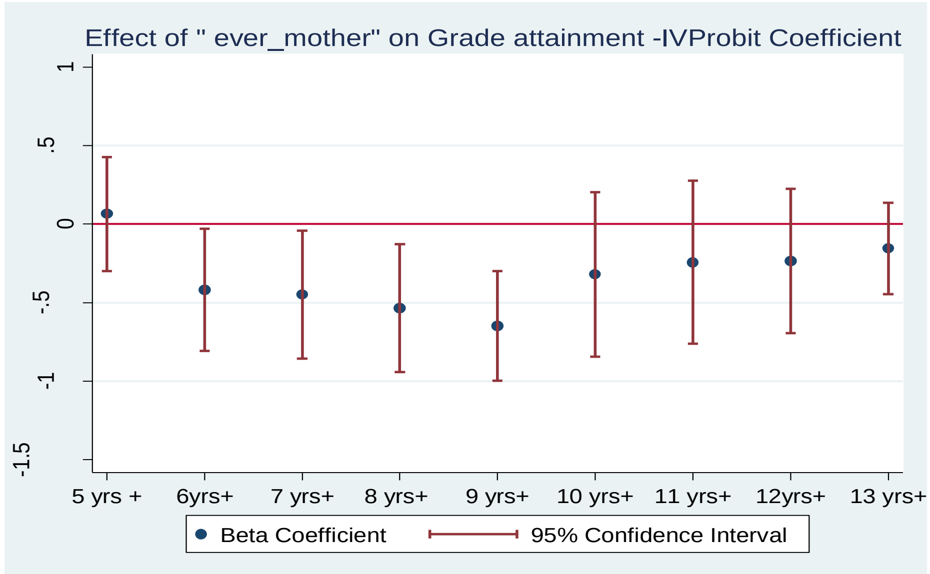
		(1) OLS	(2) IV- 2sls Access to condoms	(3) IV Exposure to Condoms /a
Panel A : Dependent Variable				
Current Enrolled	Ever _mother	-0.275***	-0.428**	-0.427***
		[0.0270]	[0.189]	(0.126)
	Fstat		11.36	11.84
	N	750	750	750
Panel B : Dependent Variable				
Years of Education	Ever _Mother	-2.029***	-2.172	-2.400
		[0.201]	[1.460]	[1.487]
	Fstat		11.36	11.84
	N	750	750	750
Panel C : Dependent Variable				
Completed 9 Years of school	Ever _Mother	-0.259***	-0.486**	-0.445***
(Lower Secondary School)		[0.0326]	[0.243]	(0.055)
	Fstat		11.36	11.84
	N	750	750	750

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses.
(a) Models with IV-Exposure to condoms in Panel A and Panel C are estimated with IV-probits. For these models, the Ever Mother coefficient is the average marginal effect and standard errors are calculated with the delta method. All the models include age cohort dummies, parents education, 2004 asset index, and extensive social infrastructure variables at the community level as well regional dummies

[53] Ranchod et al (2011) defined dropout as a dummy variable that takes value of 1 if the girl has not been enrolled at any point before completion of high school.

Table 2.8 also indicates that adolescent motherhood decreases by 2 the years of education under the OLS model and between 2.1 and 2.4 under both IV specifications; however, these later point estimates are not statistically significant. The magnitude of this effect is larger than the one found by Arceo et al. (2012) in Mexico. Using propensity score matching, the authors find that, in the long run, teenage pregnancy decreases by 1.2 the years of schooling. As we described earlier, the difference in the distribution of years of schooling between 'non- mothers' and 'ever mothers' is larger when girls are going through the secondary school cycle. This difference is supported by our empirical models. Having a child decreases the probability of completing lower secondary school by 25% under the OLS estimation, 44% when 'exposure to condoms' is used as an IV, and 48% when the IV is 'access to condoms'. The two IV point estimates are statistically significant at the 5% and 1% level, respectively. Compared with the sample mean, the IV estimate of the marginal effect of early childbearing implies a decrease from 50% to 28% in the completion of lower secondary school. We plot in Figure 2.4 the impact of early childbearing during the progression through secondary school, i.e., from having five years or more of schooling to having 13 years or more.[54] In the graph, we observe that the most adverse effect of early childbearing occurs when the young women are in the lower secondary school cycle (i.e., having 7 to 9 years or more of education). This effect of childbearing is attenuated in the upper cycle of the secondary school.

[54] These results correspond to the IV probit models but similar results are obtained from the models using access to condoms.



Notes: Beta Coefficient corresponds to the average marginal effect of the IV-probit models. Models include individual, household and community control variables

Figure 2.4 Effect of Ever Mother on Grade Attainment-IVProbits

We observe in table 2.8 that the OLS estimations underestimate the effect of the teenage pregnancy on school outcomes. If the OLS estimates have a causal interpretation, IV and OLS do not estimate the same parameter. In particular, if the response to treatment (in this case, teen fertility) is heterogeneous, then OLS captures a variance-weighted response while IV captures the response for those young women whose treatment status was affected by the instrument; i.e local average treatment effect-LATE- (Angrist and Evans, 2000).[55] Our findings are consistent with Keplinger et al. (1999) who, using a large set of variables on the costs to control fertility (i.e., contraception prevalence, abortion rates etc.) in the US, also find that the IV estimates

[55] Our sample is small enough to estimate heterogeneous effects and characterize which subgroups of the sample are being more affected by access and exposure to condoms.

of teenage fertility on educational attainment and labor market outcomes are larger than the OLS estimates context.[56] Their result differs from other studies in the US that have found that the OLS results overestimate the effect of teenage pregnancy.[57] Indeed, in light of Keplinger et al (1999), our interpretation of the higher estimates is that they reflect the marginal impact of early childbearing on schooling outcomes for that portion of the sample of young women whose fertility decisions have been affected by the variation in the access (or exposure) to condoms. Larger IV estimates might suggest that those young mothers on the margin; i.e. those girls who face higher costs of condom access and who would have avoided early childbearing had these costs been lower, experience larger human capital losses than the average young mother.[58] The bias between the OLS and the IV might be determined by the heterogeneity of the unobserved costs and benefits of the treatment (Ebestein, 2009).[59] It is possible that girls who have higher opportunity costs of dropping out from school, i.e., they are more "able and/or motivated" to keep studying, at the same time, might be more likely to engage in casual sex, and therefore, more likely to use condoms. These girls' costs of early childbearing are between those who will never have a child (never takers) and those who will always have a child (always takers). The relative

[56] Using as an IV the reforms of abortion in 1970, Angrist & Evans (2000) also find that the IV estimates of the impact of teenage pregnancy on school outcomes is larger than OLS estimates among the black young women.

[57] See for example Fletcher and Wolf (2009).

[58] This explanation is similar to the credit constraints argument that the empirical evidence has used to explain why the IV estimations of the returns to schooling are larger than the OLS (Card, 2001).

[59] Ebenstein (2009), using a sex-preference instrument for fertility, shows that the same IV has different results on labor force participation depending on the context. He shows that in the US, the OLS overestimates the IV parameter while in Taiwan the opposite happens. The author uses a conceptual framework to show that the difference in results is due to variation in the unobservable heterogeneity of benefits and costs; for example, in Taiwan, sex preferences are stronger than the US.

importance of always takers and never takers is unclear and in theory the LATE can either overestimate or underestimate the average parameter (Ebenstein 2009).

2.4.1 Impact on Cognitive Skills

We explore whether the pregnancy-related school dropout has an impact on young women's cognitive skills, measured by French and Math test scores in 2012. Table 2.9 shows that, under the OLS specification, early childbearing is associated with a loss in the order of 0.37 and 0.43 in the Math and French standardized test scores, respectively. These OLS estimates are statistically significant at the 1% level. Once we account for the endogeneity and instrument fertility with "access to condoms", this effect increases to 1.13 and 1.142, respectively, for Math and French at the 5% of statistically significance level. Using "exposure to condoms" as an IV, adolescent motherhood decreases by 1.49 and 1.56 the standardized test scores of Math and French, respectively, at the 1% of statistically significance level. The difference between the OLS and IV results are statistically different indicating that the endogeneity does have a considerable effect on the magnitude of the adolescent pregnancy impact on cognitive ability.[60]

[60] We reject the exogeneity null hypothesis with Hausman and Durbin Watson tests at the 5% significance level.

Table 2.9 : Impact of Early Childbearing on Cognitive Skills

		(1)	(2)	(3)
		OLS	IV- 2sls Access to condoms	IV-2SLS Exposure to Condoms
Panel A : Dependent Variable				
Standardized Math Score	Ever _mother	-0.371***	-1.136**	-1.495***
		[0.0637]	[0.532]	[0.570]
	Fstat		12.37	12.269
	R- Square	0.414	0.2789	0.121
	N	688	688	688
Panel A : Dependent Variable				
Standardized Score French	ever _mother	-0.429***	-1.142**	-1.569***
		[0.0611]	[0.515]	[0.567]
	Fstat		12.83	12.11
	R- Square	0.479	0.361	0.178
	N	679	679	679

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses.
Standardized test scores are calculated by subtracting the mean and dividing by the standard deviation.
All the models include age cohort dummies, parents education, 2004 asset index, and extensive social infrastructure variables at the community level as well as regional dummies

This loss in girls' cognitive ability due to pregnancy plausibly depends on how long girls have been in school. In fact, there is empirical evidence in Kenya suggesting that completing secondary school has substantial impacts on vocabulary and reasoning tests in adulthood (Ozier, 2011). We estimate OLS models of the effect of highest grade attained on the standardized test scores of Math and French using the entire cohort sample, men and women aged 21 to 23, and controlling for the same individual, household and community characteristics used in the earlier IV estimations. We are aware of the potential endogeneity of the school attainment and cognitive skills, given that there might be some unobservables that affect simultaneously the grade completed and the cognitive ability such as parental preferences. [61] Nevertheless, we do this exercise to compare the magnitude of the average effect of school attainment (highest grade attained) on the standardized test scores with our estimates of early pregnancy. Having completed lower secondary (i.e., 9 years or more of schooling) increases the standardized test scores of Math and French between 0.9 and 1.25 among young men and women in the sample (see Table A.2.3 in appendix). The longer the stay in school the larger the effect: having completed upper secondary school increases by 1.5 the standardized test scores of Math and French. This association of school attainment and test scores is statistically significant at the 1% level. Similar results are found when we estimate these OLS models only in the male sample. These point estimates are in the order of magnitude of the former IV results suggesting that the effect of early childbearing on cognitive skills is capturing the shorter stay in school due to pregnancy. Indeed, this hypothesis is consistent with prior results of Glick et al. (2009) that show a strong correlation of school attainment and test scores in Madagascar using the 2004 round of our survey.[62]

We also explore whether the effect of adolescent motherhood on the 2012 tests scores is related to earlier test score performance. We do so because it is possible that girls with previous lower scores are less motivated to stay in school and decide to get pregnant. In order to test this hypothesis, following the framework of value-added

[61] Glick and Shan (2009) do not reject the exogeneity of school attainment in a similar data of children 14 -17 years old in Senegal. Also, Glick et al (2009) found similar effects of grade attainment on test scores in different specifications including the selection into schools suggesting that this endogeneity bias should not be a concern.

[62] Glick et al (2009) find that, among children aged 14 -16 years old, attending (or have attended) grade 8 to 9 increases by 0.8 to 1.4 the standardized test scores of written Math in an OLS model and a school fixed effect model, respectively.

models for tests scores (Todd et al. 2003), we estimate the effect of the 2004 scores on the 2012 scores with and without the variable "*ever_mother"* controlling for the same set of independent covariates used in the IV estimations. Given that we only have 2004 test scores data for half of the women in the sample, we could not use "access or exposure to condoms" as an instrument of fertility. The results of this exercise are shown in Table A.2.4 of the appendix. We observe that the effect of the 2004 standardized tests scores on the 2012 standardized test scores of Math and French do not change significantly when including *"ever_mother"*.[63] Although we acknowledge the limitation in addressing the endogeniety of fertility in these estimations, the results suggest that early childbearing has an impact on the 2012 test scores of Math and French independent of the 2004 performance, at least in this portion of the women's sample.

2.4.2 Survival Analysis Results

As explained in the empirical methodology section, the Weibull hazard model allows us to calculate a predicted mean age of first birth for all the school girls who drop out after age 13.[64] Table 2.10 summarizes the effect of this "predicted age of first birth" on the school outcomes previously analyzed in the IV models. We observe that delaying by 1 year the first birth increases the probability of current enrollment by 5.6% and the probability of completing lower secondary school by 8.4%. Regarding the test scores,

[63] The marginal increase in the R^2 of the regression that includes "ever mother" implies that the fertility variable adds information to the value-added models
[64] Similar results are obtained when we use predicted median age of first birth.

postponing by 1 year the first birth increases by 0.19 and 0.21 the standardized test scores of Math and French, respectively.

Table 2.10 Effect of Predicted Age of First Birth on School Outcomes

	Current Enrollment	9 or more Years of Schooling	Z-Score French	Z-Score Math
	(1)	(2)	(3)	(4)
Predicted Age First Birth (Mean)	0.056** [0.026]	0.084** [0.030]	0.190*** [0.0614]	0.211*** [0.0642]
N	750	750	688	679

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors
Age of First Birth was predicted after estimation of Weibull models in the first stage
Models (1) and (2) are estimated with probit models. Coefficients are average marginal effects All the models include the individual, household and community control variables

These findings of the survival models are consistent with our two-stage models results. Consider an average girl of our sample who gets pregnant in school and has accumulated 7 years of education. If she has the option of postponing her first birth by at least 5 years, under the assumption of no grade repetition, she would be 40% likelier to complete at least lower secondary school. Similarly, if she can improve by 0.2 the standardized tests scores of Math and French each year of school attainment after 5 years of delaying the childbirth, she would have a return of around 1 standard deviation in her tests scores. This result is a very close estimate to the effect of "ever mother" on the cognitive skills using the IV models presented earlier.

The paper results allow us to establish a detrimental and causal effect of teenage pregnancy on young women's human capital in Madagascar, deviating from the findings of Azevedo et al. (2012) in Mexico and Ranchod et al. (2011) in South

Africa. We need to bear in mind that the estimation of fertility impacts on socioeconomic outcomes depends on the identifying instrument employed, since there is heterogeneity in the individual responses to the specific chosen instrument. In other words, the effect estimated from variation in a policy variable represents a *specific LATE* of modifying the fertility of certain groups in the population (Shultz, 2007).[65] Therefore, our IV results are limited to the sample of young girls whose childbearing decisions are induced by the access (exposure) to condoms which probably is not representative of the average young school girl in Madagascar; however, these girls have higher opportunity costs when getting pregnant.

2.4.3 Robustness Checks

To validate our hypothesis that access to condoms should only affect young women's schooling outcomes through the avoidance of their first birth and not through other alternative channels; we estimate the reduced form of access to condoms on the school outcomes of young men in the same age cohort. If this placebo test is valid, we should expect that the direct effect of access to condoms on male's education outcomes is not statistically significant.

[65] This limitation of the IV estimation is also common to studies that use natural experiments such as miscarriages to identify the effects of teenage pregnancy. By comparing teenage mothers to those girls who have had a miscarriage, causal effect concerns only the atypical subsample of the relevant population (Keplinger et al, 1999).

Table 2.11 Reduced form of Access to Condoms on Male and Female School Outcomes

Dependent Variables		Current Enrollment	Years of Schooling	Completed Lower Secondary	Z-Score French	Z-Score Math
Panel A:						
Outcomes for Young Men	Access to Condoms	-0.00512	0.0567	0.0352	0.121	0.0863
		[0.0341]	[0.475]	[0.0676]	[0.151]	[0.141]
	Adj -R^2	0.129	0.371	0.379	0.311	0.250
	N	723	723	723	664	675
Panel B :						
Outcomes for Young Women	Access to Condoms	0.0765**	0.388	0.0869*	0.233**	0.224**
		[0.0332]	[0.281]	[0.0448]	[0.0959]	[0.0920]
	Adj -R^2	0.162	0.436	0.420	0.363	0.286
	N	750	750	679	688	750

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively.
Robust Standard errors reported in parentheses.
All the models include the individual, household and community control variables

We construct a similar sample of young men aged 22 on average and who drop out of school after age 13. In this sample, we control for the same set of individual, household and community characteristics included in the IV models. Table 2.11 shows that the effect of 'access to condoms' on the young men's current enrollment, years of education, completion of lower secondary and test scores. Compared to the same reduced form for young women, the point estimate of access to condoms on the male schooling outcomes are much smaller in magnitude and not statistically significant.

In addition to the condom program placement checks that we present earlier in the empirical section, we also estimate our IV models controlling for the two different measures of 2006 fertility at the community level: number of births and number of women who died during or after child delivery. Table 2.12 shows the OLS and IV models for the current enrollment and the standardized scores of Math and French outcomes controlling for these 2006 fertility variables and the rest of individual, household and community covariates.[66] The effect of 'ever mother' on these school outcomes is robust to the inclusion of the 2006 fertility variables supporting our finding that there is no evidence for any non-random program placement.

[66] These models were also estimated for years of schooling and completion of lower secondary school and the results are robust to the 2006 fertility controls.

Table 2.12 Robustness Check: OLS and IV Models with 2006 Fertility Variables

Panel A. OLS and IV Models Controlling for 2006 Number of Births at the Community Level

	Current Enrollment		Standardized Score of Math		Standardized score of French	
	OLS	IV -Access to condoms	OLS	IV -Access to condoms	OLS	IV -Access to condoms
Ever Mother	-0.275***	-0.443**	-0.371***	-1.077**	-0.428***	-1.205**
	[0.0270]	[0.186]	[0.0637]	[0.492]	[0.0611]	[0.487]
2006 No. of Births	-0.00000270	-0.00000448	0.0000396	0.0000358	0.0000256	0.0000195
	[0.0000150]	[0.0000149]	[0.0000286]	[0.0000308]	[0.0000288]	[0.0000335]
Missing No. of births	0.00777	0.00977	-0.0311	-0.0251	0.0464	0.0610
	[0.0619]	[0.0614]	[0.122]	[0.136]	[0.113]	[0.131]
N	750	750	688	688	679	679

Panel B OLS and IV Models Controlling for 2006 Maternal Mortality at the Community Level

	Current Enrollment		Standardized Score of Math		Standardized Score of French	
	OLS	IV -Access to condoms	OLS	IV -Access to condoms	OLS	IV -Access to condoms
Ever Mother	-0.274***	-0.370**	-0.372***	-0.998**	-0.430***	-1.183**
	[0.0270]	[0.188]	[0.0634]	[0.500]	[0.0611]	[0.512]
2006 No. of Women dead during/after child delivery	0.000864	0.000347	-0.0110	-0.0136*	-0.00332	-0.00721
	[0.00350]	[0.00362]	[0.00686]	[0.00782]	[0.00693]	[0.00847]
Missing	-0.0529	-0.0497	-0.199*	-0.174	-0.0268	0.00635
	[0.0462]	[0.0467]	[0.111]	[0.116]	[0.111]	[0.121]
N	750	750	688	688	679	679

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses.
Standardized test scores are calculated by subtracting the mean and dividing by the standard deviation.
All the models include age cohort dummies, parents education, 2004 asset index, and extensive social infrastructure variables at the community level as well as regional dummies.

2.5 Conclusions

Empirical evidence on the economic consequences of adolescent pregnancy is scarce in developing countries, particularly, in Sub-Saharan Africa. We contribute to this gap in the literature by addressing whether early childbearing affects school dropout and cognitive skills among young women in Madagascar, a low-income and high-fertility country. Using a panel data survey combined with community censuses, we address the endogeneity between fertility and education decisions by instrumenting the young woman's access to condoms at the community level, and her exposure to condoms since she was 15 years old. We control for a complete set of covariates at the community level to account for the potential endogeneity of program placement.

Our findings point out a detrimental and causal effect of teenage pregnancy on young women's human capital in Madagascar. Our IV results indicate that young women's early childbearing increases by 42% their likelihood of dropping out of school and decreases by 44% their chances of completing lower secondary school (i.e., 9 years of more of schooling). These findings suggest that early pregnancy and schooling are mutually exclusive in Madagascar.

Furthermore, this school-pregnancy related dropout is associated with a decrease in the standardized test scores of Math and French in the order of 1.1 to 1.5 standard deviations. This magnitude is comparable to the effect of secondary school attainment on the test scores suggesting that the shortened stay in school due to pregnancy has detrimental effects on cognitive skills. These results on cognition are a unique contribution to the empirical literature in developing countries. We also obtain consistent results when we model the age of first birth using hazard models in the first stage. Delaying the first birth by a year increases the probability of current enrollment by 5%, the likelihood of completing secondary school by 8% and the test scores of Math and French by 0.2 standard deviations.

Our results underline the potential role of policies that can prevent early childbearing and those that allow teen mothers to catch up with their education, in enhancing young women's human capital investment. In particular, the results from our instrumental variable approach suggest that reproductive health and family planning policies that lower the costs of postponing the first birth among young women can have human capital gains beyond the prevention of poor pregnancy outcomes, such as risks of maternal health and low birth. This evidence is consistent with findings from a large family planning program in Colombia that enabled young women to postpone their first birth, thus allowing them to increase their years of education and labor participation in the formal sector (Miller, 2010).

More broadly, there is an ongoing debate on the effectiveness of reproductive health policies in developing countries, particularly on whether access to family planning policies reduces total fertility and improves socio-economic outcomes (Canning and Schultz, 2012). In this context, our findings suggest that, regardless of the total fertility effect, the timing of postponing the first birth is crucial to increase women's education and human capital. However, further research should analyze if this reduction in teen fertility is translated in a woman's lifetime fertility reduction as well as improvement in her and their families' economic outcomes. For instance, young women's larger human capital can also be translated into better health and education outcomes of their children, breaking potential channels of intergenerational transmission of poverty. Therefore, these policies that aim to reduce teenage pregnancy will impact not only young women's economic opportunities but those of their children.

Further research on the effectiveness of reproductive health and family planning policies is timely in Sub-Saharan Africa countries, such as Madagascar, since they are facing a demographic dividend: the number of young people aged 12 -24 is larger than ever, representing a unique opportunity to reap the benefits of enhancing young women's human capital.

REFERENCES

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Angrist, J. D., & Evans, W. N. (2000). Schooling and labor market consequences of the 1970 state abortion reforms. Research in Labor Economics, 18, 75-113

Arceo-Gomez, E., & Campos-Vázquez, R. (2012). Teenage Pregnancy in Mexico: Evolution and Consequences. *Available at SSRN 2198089*.

Azevedo P, Lopez-Calva, L Perova, E (2012). Is the baby to Blame? An Inquiry into the consequences of Early Childbearing. Policy Research Working Paper 6074. Washington, D.C

Angeles G, Guilkey DK, Mroz, TA (2005a). The effects of education and family planning programs on fertility in Indonesia. *Econ Devel Cultural Change* 54(1):165–201.

Baird, S., McIntosh, C., & Özler, B. (2011). Cash or condition? Evidence from a cash transfer experiment. *The Quarterly Journal of Economics*, qjr032.

Bertrand M., Duflo E, and Mullainathan, S (2004). How Much Should We Trust Differences-in-Differences Estimates?. *The Quarterly Journal of Economics*, 119(1): 249-275.

Canning, D., & Schultz, T. P. (2012). The economic consequences of reproductive health and family planning. *The Lancet*, *380*(9837), 165-171

Chong, A., Gonzalez-Navarro, M., Karlan, D., & Valdivia, M. (2013). Effectiveness and Spillovers of Online Sex Education: Evidence from a Randomized Evaluation in Colombian Public Schools (No. w18776). National Bureau of Economic Research.

Duflo, E., Dupas, P., & Kremer, M. (2011). Education, HIV and early fertility: Experimental evidence from Kenya. *UCLA manuscript*.

Fletcher, J. M., & Wolfe, B. L. (2009). Education and Labor Market Consequences of Teenage Childbearing Evidence Using the Timing of Pregnancy Outcomes and Community Fixed Effects. *Journal of Human Resources*, *44*(2), 303-325.

Glick, P., Handy, C., & Sahn, D. E. (2011). Schooling, Marriage, and Childbearing in Madagascar. *Cornell Food and Nutrition Policy Program Working Paper*, (241).

Glick, P., Randriamamonjy, J., & Sahn, D. E. (2009). Determinants of HIV knowledge and condom use among women in Madagascar: an analysis using matched household and community data. *African Development Review*, *21*(1), 147-179.

Grant, M . and Hallman, K. (2008). Pregnancy-related school dropout and prior school performance in KwaZulu/Natal, South Africa.*Studies in Family Planning.* 39(4): 369-382.

Hotz, V. J, Williams, S and Sanders, S (2005).Teenage childbearing and its life cycle consequences: exploiting a natural experiment. *The Journal of Human Resources*, Vol 40. No.3.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467-475.

Klepinger, D., Shelly Lundberg, and R. Plotnick, (1999). Teen Childbearing and Human Capital: Does Timing Matter? .Mimeo, University of Washington, October.

Levine, D. I., & Painter, G. (2003). The schooling costs of teenage out-of-wedlock childbearing: analysis with a within-school propensity-score-matching estimator. *Review of Economics and Statistics*, *85*(4), 884-900.

Lloyd, Cynthia and Mensch, Barbara (2008) Marriage and Childbirth in Dropping out from School: An Analysis of DHS Data from Sub-Saharan Africa. *Population Studies,* Vol. 62, No. 1

Marletto, L., Lam and Ranchond, V. (2008). Sexual Behavior, Pregnancy, and Schooling Among Young People in Urban South Africa. *Studies in Family Planning* 2008; 39[4]: 351–368).

McQueston, K., Silverman, R., & Glassman, A. (2012). Adolescent Fertility in Low- and Middle-Income Countries: Effects and Solutions. *Center for Global Development Working Paper*, *295*.

Miller, Grant (2010) “Contraception as Development? New Evidence from Family Planning in Colombia,” The Economic Journal, Vol. 120, No. 545, pp. 709–736.

Ozier, O. (2011). The impact of secondary schooling in Kenya: a regression discontinuity analysis. Berkeley manuscript.

Population Reference Bureau (2012) “Improving the Reproductive Health of Sub-Saharan Africa’s Youth: A Route to achieve the Millennium development goal” Available at: www.prb.org/pdf10/youthchartbook.pdf

Pörtner C., Beegle K, Christiaensen L (2012). Family planning and fertility: Estimating program effects using cross-sectional data. Working papers. Washington State University

Ranchhod, V., Lam, D., Leibbrandt, Z., and Marteleto, L. (2011). Estimating the effect of adolescent fertility on educational attainment in Cape Town using a propensity score weighted regression. Southern Africa Labor and Development Unit, Working paper No 59

Ribar , D (1994).Teenage Fertility and High School Completion. *The Review of Economics and Statistics*, Vol. 76, No. 3.

Rogers, William H. (1993). "Sg17: Regression standard errors in clustered samples." Stata Technical Bulletin 13: 19-23.

Thomas, Duncan. (1999) "Fertility, Education and Resources in South Africa," in Critical Perspectives on Schooling and Fertility in the Developing World. C. Bledsoe et al. eds. Washington, DC: National Academy Press.

Sharp, M., and Kruse, I. (2011). Health, Nutrition, and Population in Madagascar, 2000-09 (No. 216). World Bank-free PDF.

Schultz, T. P. (2007). Population policies, fertility, women's human capital, and child quality. *Handbook of Development Economics*, *4*, 3249-3303.

United Nations Population Fund. (2013). Motherhood in Childhood: Facing the Challenge of Adolescent Pregnancy. *State of World Population 2013*. New York: United Nations.

World Bank (2013) World Development Indicators

APPENDIX

Table A.2.1 First Stage using Clustered standard Errors

	"ever_mother"	"ever_mother"
Condom Access	**-0.179***** **[0.0626]**	
Condom Exposure_15 yrs		**-0.0234***** **[0.00864]**
Asset Index 2004	-0.0122 [0.0270]	-0.0135 [0.0267]
Mother is alive (Y=1)	-0.0390 [0.0586]	-0.0408 [0.0588]
Father is alive (Y=1)	-0.0188 [0.0414]	-0.0219 [0.0404]
Mother's years of education	-0.0147*** [0.00544]	-0.0147*** [0.00554]
Father's years of education	0.000587 [0.00609]	0.000345 [0.00615]
Community health Center (CSB2)	0.0808 [0.0998]	0.0745 [0.101]
Community Hospital (CHD1)	-0.000210 [0.0556]	-0.00206 [0.0564]
Upper Secondary (y=1)	-0.0965 [0.0621]	-0.0903 [0.0653]
Piped Water (Y=1)	0.0951* [0.0498]	0.0854* [0.0462]
Access to weekly market (Y=1)	-0.104** [0.0468]	-0.112** [0.0454]
Access to paved road all year(Y=1)	0.0683 [0.0491]	0.0607 [0.0491]
Electricity(Y=1)	0.0597 [0.0544]	0.0509 [0.0549]
Upper Secondary at 10 yrs old	0.0486 [0.0696]	0.0559 [0.0712]
CSB2 at 10 yrs old	-0.0276 [0.0666]	-0.0295 [0.0659]
Electricity at 10 yrs old	-0.222*** [0.0826]	-0.223*** [0.0741]
Remotness index 2001	-0.00527 [0.0186]	-0.00277 [0.0178]
Urban (Y=1)	0.0135 [0.112]	0.0251 [0.103]
N	750	750
First Stage	**8.1618**	**7.3442**
R-sq	0.136	0.1365

Notes: * p<0.10, ** p<0.05, *** p<0.01" Robust clustered standard errors in brackets. Models include cohort age and regional dummies not shown

Table A.2.2 IV Results of the effect of Ever Mother on 2012 School Outcomes Clustering Standard Errors

	(1)	(2)	(3)
2012 School Outcomes	**OLS**	**IV- Access to condoms**	**IV -Exposure to condoms**
Current Enrollment	-0.275***	-0.428**	-0.427***
std error	[0.0277]	[0.208]	(0.126)
Compl. Lower Second School	-0.259***	-0.486*	-0.445***
std error	[0.0336]	[0.272]	(0.064)
Years of Schooling	-2.029***	-2.172	-2.400
std error	[0.193]	[1.925]	[1.786]
Zscore Math	-0.371***	-1.136*	-1.495**
std error	[0.0722]	[0.650]	[0.709]
Zscore French	-0.429***	-1.142	-1.569**
std error	[0.0765]	[0.745]	[0.755]

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively.

Clustered Robust Standard errors reported in parentheses.

All the models include the individual, household and community control variables

In Column 3, binary outcomes are estimated using IV-probit models

Table A.2.3 School Attainment and 2012 Standardized scores of French and Math

	All Sample		Male	
	Z-Score Math	**Z-Score French**	**Z-Score Math**	**Z-Score French**
Complete Primary (5 yrs sch)	0.345***	0.306***	0.351***	0.289***
	[0.0616]	[0.0584]	[0.0839]	[0.0797]
Some College (6-8 yrs-sch)	0.853***	0.769***	0.952***	0.905***
	[0.0649]	[0.0608]	[0.0895]	[0.0830]
Complete College (9 yrs sch)	0.987***	0.984***	0.979***	1.066***
	[0.0656]	[0.0616]	[0.0864]	[0.0809]
Some Lycee (10-11 yrs sch)	1.254***	1.353***	1.262***	1.421***
	[0.0756]	[0.0694]	[0.0995]	[0.0936]
Complete Lycee (12 yrs sch)	1.574***	1.620***	1.637***	1.715***
	[0.0691]	[0.0624]	[0.0928]	[0.0802]
Superior (12 and more)	1.964***	1.956***	2.131***	2.042***
	[0.0845]	[0.0738]	[0.127]	[0.105]
N	1363	1343	675	664
adj. R-sq	0.574	0.658	0.592	0.679

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively.
Robust Standard errors reported in parentheses. Individuals included in the sample dropout from school at age older than 13. All the models include the individual, household and community control variables

Table A.2.4 2004 Test Scores Effects on 2012 Test scores of Math and French

	2012 Z-Score Math		2012 Z-Score French	
	(1)	(2)	(1)	(2)
2004 Z-score Math	0.190*** [0.0491]	0.186*** [0.0525]		
2004 Z-Score French			0.222*** [0.0494]	0.206*** [0.0474]
Ever_mother		-0.309*** [0.0847]		-0.329*** [0.0794]
N	402	402	390	390
Adj. R^2	0.381	0.401	0.455	0.479

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses. Women included in the sample dropout from school at age older than 13. 2012 and 2004 z-scores are the standardized scores with mean 0 and standard deviation of 1. All the models include the individual, household and community control variables

Table A.2.5 Effect of Predicted Age of First Birth on School Outcomes Changing origin at 12 years old

Panel A: First Stage Age of First Birth

	Hazard ratio	Robust Standard error	z	P>|z|
Condom Exposure 15 yrs	0.940	0.018	-3.29	0.001

Notes: p =2.58 (std error 0.117) ; No of observations 750; Wald Chi2= 114
All the models include the individual, household and community control variables

Panel B: Second Stage School Outcomes

	Current Enrollment	9 or more Years of Schooling	Z-Score French	Z-Score Math
	(1)	(2)	(3)	(4)
Predicted Age First Birth (Mean)	0.043** [0.02]	0.064*** [0.024]	0.127*** [0.0476]	0.156*** [0.0489]
N	750	750	688	679

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust standard errors. Models (1) and (2) are estimated with probit models, thus coefficients are average marginal effects and standard errors calculated by delta method .All the models include the individual, household and community control variables

CHAPTER 3

THE ROLE OF MATERNAL COGNITVE SKILLS IN CHILDREN'S HEALTH IN MADAGASCAR

Abstract

Chronic malnutrition affects half of the children under age of five in Madagascar and little is known about the role of mother's human capital in improving child health. This paper analyzes the effect of maternal cognitive skills, measured by the sum of her Math, French and Life Skills test scores, on children's height-for-age in Madagascar. Using a panel survey, we address the potential endogeneity between maternal schooling and child health by adding to the OLS models an extensive set of mother's childhood characteristics, her personality traits, as well as community variables where mother and child live. Our findings show that the relationship between mother's years of schooling and child health is mediated by her cognitive skills. One standard deviation in the mother's standardized total score increases her child's height-for-age between 0.33 and 0.47 standard deviations. This effect of maternal cognitive ability on child health is not channeled through an increase of household wealth or use of available health infrastructure. These findings are robust to the inclusion of the mother's non-cognitive skills, measured by her personality traits.

3.1 Introduction

Maternal education is positively associated with children's health (Strauss and Thomas, 1995). This association is particularly important given that early life investments in human capital can define later economic outcomes (Heckman, 2006). In fact, improvements in nutritional status during early childhood can positively affect adult economic productivity in developing countries (Hoddinott et al., 2008). Therefore, the deleterious impact of a women's lower education not only has a detrimental impact over their life course, but also potentially contributes to the intergenerational transmission of poverty. Unlike empirical evidence on the returns of mothers' years of schooling on children's health, there are few studies that analyze maternal human capital directly as measured in terms of cognitive skills; and this is especially the case in the context of Sub-Saharan African countries.

This paper analyzes the role of maternal cognitive skills, measured by the standardized sum of the mother's Math, French and Life Skills test scores, on their children's height-for-age in Madagascar. Malnutrition is widespread in Madagascar, 50% of children younger than 5 years old are stunted and 35% are severely stunted (2009 Demographic Health Survey). Indeed, Madagascar has the sixth highest rate of malnutrition in the world and ranks higher than its income country peers in Sub Sahara Africa (Sharp, M. and Kruse, I. 2011).

To examine the effect of maternal cognitive skills on children's health, we use a panel data survey designed to capture the transition from adolescence to adulthood of a cohort of young people in Madagascar, interviewed the first time in 2004 and then re-interviewed in 2012. Specifically, our sample of children under age of 5 includes those children born before 2012 to the young women first interviewed in 2004. The average age of the young mothers is 22 and they have completed six years of education on average. Only 7.7% of them have completed secondary school. This 2012 data include mother's cognitive and non-cognitive skills measures. In particular, the mother's test scores in Math, French and Life Skills are designed to assess their knowledge accumulated in and out of school. Our primary interest is in determining the relative role of these cognitive skills, which only in part reflects years of formal education. Indeed, while schooling has been shown to affect cognitive skills in Madagascar and elsewhere, so do other factors, specifically, local school quality, family characteristics such as wealth and parental education and community characteristics.

Female education can affect children's health through several channels. First, more educated mothers might have higher wages in the labor market increasing the economic resources in the household. Second, more educated mothers can have more bargaining power in household decision making allocating more resources to child human capital investment. Third, more educated women can have better health knowledge and more access to new information.[67]

Unlike schooling attainment, the role of maternal cognitive skills on child health has been scarcely explored. Using a longitudinal panel in Guatemala, Behrman et al. (2009) show that cognitive skills have a larger impact than years of education on child's health. The authors use Raven and reading test scores to measure mothers'

[67] See Rubacalva et al (2004) and Glewe (1999)

cognitive ability. Similarly, using Raven tests as a maternal cognitive measure, Rubacalva et al. (2004) analyze the channels through which mother's cognitive ability affects child height in Mexico. The authors show that maternal cognitive ability is related to the capacity of child rearing for the first time; i.e., the mother's Raven test effect is more important for the first born than non-first born children. Also, mother's cognitive ability reflects her capacity to take advantage of the community resources and to increase household wealth. The effect of maternal cognitive ability on child health is not related to her childhood background such mothers' parents' education and type of school attended. In Morocco, Glewwe (1999) shows that health knowledge is a critical determinant of child health and while such skills are obtained outside the classroom literacy and numeracy are the essential tools for learning these skills. Indeed, Alderman and Christiansen (2004) in Ethiopia also find that maternal nutrition knowledge is an important determinant of child height-for-age. Along the same lines, Block (2007) analyzes the effect of maternal schooling and nutrition knowledge in the child micronutrient intake finding that the effect of maternal education is partially mediated through nutrition knowledge and household expenditure in Indonesia.

One major concern in examining the impact of cognitive ability on child health is the potential endogeneity between maternal education and child health due to unobserved characteristics affecting both outcomes. One alternative to address this issue is to implement an instrumental variable (IV) approach. This is extremely challenging in terms of identifying strong instruments that strongly affect mother's schooling while at the same time not having a plausible direct pathway affect children's health. Instead, we address the endogeneity issue by gradually introducing, in the OLS models of the child health on maternal cognitive skills, a rather exhaustive set of covariates that might affect both variables and will greatly mitigate, although never fully remove all possible concern over the endogeneity of mother's cognition and child health. Specifically, we control for mothers' years of schooling and their height, as well as an extensive set of mother's childhood socioeconomic background such as her parents', grandparents' and siblings' education as well as the primary school quality indexes of the community where the mothers lived in 2004 when they were 13 years old on average. We also include a 2004 asset index and the mother's spouse education as proxy to control for income resources.

To address the child's health heterogeneity at the community level, we include 2012 community variables such as access to, or the existence of, health centers, piped water and electricity. In particular, we include a 2001 remoteness index that help us control for unobservables related to economic development at the community level, ten years prior to the time of the survey (2012). Additionally, we present the models with community fixed effects as a robustness check. Using a unique feature of the data, we are able to test whether the gradient of maternal cognitive ability on child health is affected by mothers' non-cognitive skills, measured by her personality traits: openness, conscientiousness, extraversion, agreeableness and neuroticism.

Our findings show that maternal cognitive skills have a positive and statistically significant return on children's health. In our preferred specifications, one standard deviation in the mother's total score increases the child height-for-age by 0.33-0.47 standard deviations. This result is robust to the inclusion of other mother's human capital measures such as years of schooling and height as well as her socioeconomic childhood characteristics. The results are also robust in the community fixed effects models. The lack of an independent effect of school attainment is noteworthy, suggesting that the role of formal education on child health is captured by the mother's cognitive skills. We also find that the cognitive skills' effect on child health does not reflect a wealth effect; the result is robust to the inclusion of a 2004 asset index and spouse education, indicators of household economic resources. We also find that the interaction of mother's standardized total score and the access to health center is not statistically significant, indicating that maternal cognitive ability is not related to a better use of health infrastructure at the community level. Additionally, our results suggest that the effect of mother's cognitive ability on child height-for-age is independent of her non-cognitive skills, measured by her personality traits.

Our paper thus contributes to the limited empirical evidence of the effect of maternal cognitive skills on child health status in Sub-Saharan Africa, particularly in Madagascar where this is the first study to provide such evidence. Our paper analyses the effect of mother's cognitive ability on child's long term health using Math, French and Life Skills test scores designed to measure young women's knowledge in and out of school. This contrasts with studies that have used nonverbal and innate measures of cognition such as Raven Tests. We contribute to former studies that have analyzed the effect of maternal numeracy and literacy skills on child health, by adding as control variables an unusual set mother's educational background variables such as education of her parents, siblings and grandparents as well as economic and community environment characteristics in which the mother grew up such as quality indexes of school attended. In fact, to our knowledge, this is the first study that analyzes the role of non-cognitive skills in the gradient of maternal education on child health in the context of developing countries.

This paper is organized as follows. Section 2 describes the panel data set and the context of Madagascar. Section 3 presents the empirical methodology to estimate the effect of the cognitive skills on children's health. Section 4 discusses the results, while the last section concludes and discusses policy implications.

3.2 Data and Descriptive Statistics

The Madagascar Life Course Transition of Young Adults Survey (MLCT) 2011-12 re-interviewed a cohort of young adults, who were 21-23 years old at the time of the survey, and gathered detailed information on them, their spouses, and the households in which they resided at the time of survey, as well as their communities. Additionally, the survey collected detailed retrospective event histories on the cohort members, their families and the range of economic and life-course events and experiences going back to 2004, the year in which the cohort members were last surveyed in the *Enquête sur la Progression Scolaire à Madagascar* (EPSPAM) survey.

The MLCT data contain direct and indirect measures of human capital, including tests of cognitive skills in Math, French and Life Skills and a set of questions to measure personality traits, sometimes referred to as non-cognitive skills. The 2011-12 survey also updates histories of community level services and economic and social infrastructure. We also rely on the 2004 survey of the primary and secondary schools in the community in which the women grew up, which contains extensive information that allow us to construct school quality and infrastructure indexes.[68] Additionally, we complement the information at the community level with two other independent censuses of all the villages in Madagascar conducted in 2001 and 2007 including information on the basic public services and infrastructure.

This paper uses the sample of children who are sons and daughters of the female cohort members, young mothers on average 22 years old in 2012. We measure child health using the height-for-age Z-Score (HAZ)[69], a good indicator of child chronic malnutrition and general health status.[70] Additionally, the anthropometric indicators have the advantage that their errors in measurements are not correlated with other unobservable variables (Federov and Sahn, 2005). We restrict our analysis to the sample of 500 children aged 0 to 60 months since the differences in height (standardized by age and gender) within this age range are largely determined by differences in the socioeconomic environment in which they are raised, rather than ethnic or racial differences.[71] The average HAZ is -1.61, 47% of the children under age of 5 are stunted and 28% are severely stunted. These figures are consistent with 2009 Demographic Health Survey in Madagascar. [72]

---

[68] The school quality variables include distance from the center of the community, number of blackboards per student, number of textbooks per student, availability of electricity, number of toilets, as well as principal and teachers' years of education and experience.

[69] The height-for-age is a standardized measure of child's height by age and sex. Standardization is done by fitting a standard normal distribution to the growth curves of a healthy population, designed by the WHO, using an age and gender specific heights.

[70] We restrict our child's health measure to the height-for-age indicator because the child's weight and birth weight have measurement error. The last measure was reported by the mothers.

[71] Medical literature has shown that standards for child growth during prepubescent age are common across racial and ethnic groups (WHO, 1983)

[72] Stunting is defined as height-for-age z-score less than -2 and severe stunting is defined as height-for-age z-score less than –3.

Table 3.1 summarizes the variables used in the analysis. There is missing information in the mother's test scores variables: 12% and 13% in the Math and French tests, respectively and 3% for the Life skills. Since the total score is constructed as a sum of the three scores, the missing information is 3%. To avoid losing this portion in the sample estimation, we create a dummy variable to indicate that the total score was missing; in such cases the total score is set equal to the median. We have a similar issue of missing data with the spouse education since only 70% of the young mothers report that they are married. Therefore, we also create a dummy variable to indicate that the spouse education was missing and replace the missing values with the median of spouse education.

The average education among mothers is 6 years; 30% have some primary school, 17% have completed primary school, 16% have some lower secondary school (9 years completed), and only 7.7% completed the entire secondary cycle. Also, these young mothers have lower test scores than the sample mean of the female cohort members and their counterparts who are non-mothers. Indeed, Herrera and Sahn (2013) show that this lower school attainment and cognitive skills is caused by early childbearing leading to dropping out of school.

Table 3.1 Summary of Statistics

Variable	N	Mean	Std. Dev.
Child characteristics			
Gender	500	1.482	0.500
Height-for-age	500	-1.615	2.39
Age in months	500	2.628	1.373
Mother's Characteristics			
Mother Total Standardized Score	500	0.000	0.995
Missing Total Standardized Score	500	0.036	0.186
Age	500	22.102	1.22
Mothers ln height	495	5.043	0.044
Mother's Years of Schooling	500	6.082	3.221
Grandmothers Education	498	3.863	3.341
Grandparents Education	498	4.556	3.805
Great Parents Education (Max)	500	2.248	1.170
Missing Great parents education	500	0.110	0.313
Mother's No of siblings	500	3.478	2.054
Mother's Siblings average education	500	2.821	2.364
2004 Asset Index	500	-0.125	0.792
Spouse Education	500	6.256	3.435
Missing Spouse Education	500	0.288	0.453
Mother's Non Cognitive Skills			
Openness	499	-0.305	1.008
Agreeableness	499	-0.110	1.057
Conscientiousness	499	-0.182	1.052
Extroversion -Ex	499	-0.230	1.017
Neuroticism-emsz	499	0.143	1.058
Community Characteristics			
Access to Community Health Center(Y=1)	500	0.678	0.468
Urban (Y=1)	500	0.188	0.391
Access to Piped Water (Y=1)	500	0.450	0.498
Access to Electricity (Y=1)	500	0.380	0.486
2001 Remoteness Index	500	2.574	1.361
2001 Remoteness Index quadratic	500	8.474	7.820
2004 School Facility Index	491	-0.129	0.713
2004 School Teaching Index	491	-0.079	0.457

Figures 3.1 shows the percentage of children stunted by mothers' years of education and terciles of the total standardized test score. We observe a positive gradient between child health and both mother's education and cognitive skills. Among the children of mothers in the lowest tercile, 51% are stunted. This proportion decreases to 42% for mothers in the highest tercile.

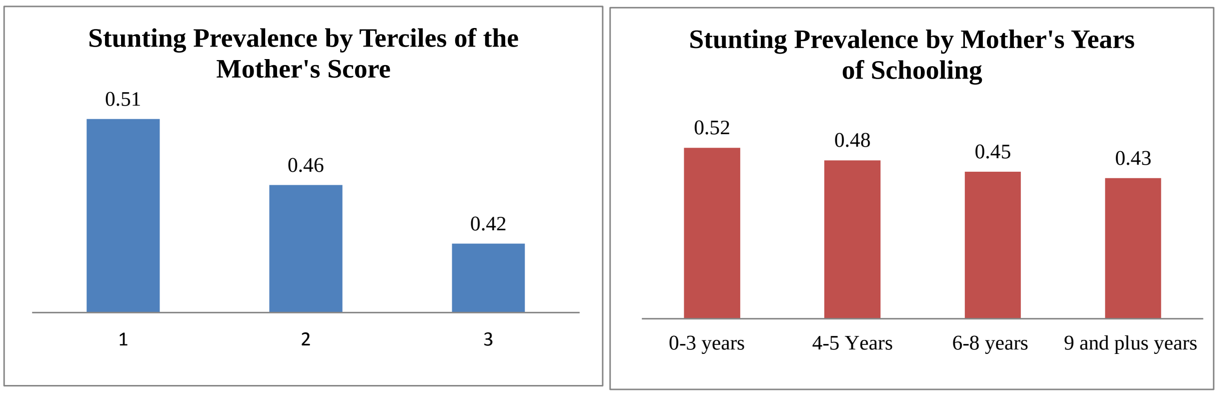


Figure 3.1 Stunting Prevalence by Mother's Total Score and Years of Schooling

We also observe higher rates of malnutrition among children whose mothers have 3 or less years of schooling compared to the rates among children whose mothers have 9 or more years of schooling. Although, statistical tests do not show a significant difference by schooling, unlike the case of the test scores where the means of mothers' test scores is significantly lower (higher) among stunted (non-stunted) children. This might be due to the fact that there is more variance in the test scores than in the years of schooling.

3.3 Empirical Methodology

The conceptual framework underlying our empirical strategy is the standard household's maximization utility function defined over leisure, child health and other home and market-purchased goods while facing a budget constraint and a biological production function. We estimate the solution to this household optimization problem, a reduced form of the demand for child health status, *Hi*, measured by height-for-age, in function of the mother's cognitive skills, mother's childhood socioeconomic background as well as community characteristics.[73]

One concern related to estimating this reduced demand for child health is the potential endogeneity between maternal cognitive skills and child health due to omitted variables, for example, unobserved characteristics of the mother or household variables that might affect simultaneously the human capital of the mother and child health. Except for few studies, most of the estimates of the effect of mother's education on children's health have neglected this potential endogeneity (Thomas et al; 1990; Christiansen and Alderman, 2004; Block, 2007). One alternative to address this issue is to implement an instrumental variable (IV) approach; nevertheless, unless there is a quasi-experimental design, it is difficult to find an exogenous variable that highly correlates with mother's schooling and does not directly affect child health.[74]

[73] Other alternative is to estimate the child health production directly but empirically this approach is challenging. It requires detailed information on different health inputs such as nutrient intake, health clinic utilization among others, that are endogenous and difficult to instrument (Christiansen and Alderman, 2004).

[74] In the context of developing countries some studies that have used quasi-experimental studies to measure the maternal education effect on child health are: Duflo et al. (2001) in Indonesia and Gunes (2014) in Turkey.

For instance, Glewwe (1999) uses the mother's parents and sisters' education as an IV for maternal education.[75] He does so despite recognizing the challenge of validating the exclusion restriction, and in particular, how child grandparents' education can be correlated with values and attitudes during childbearing that can directly affect a child's health. This is especially problematic in a context such as Madagascar where young mothers generally live in close proximity to their mothers. In fact, in our sample, 30% of the young mothers still live with their parents thus excluding the possibility of using mother's parents' education as an instrument

Alternatively to an instrumental variable approach and exploiting the extensive inter-generational information in the Malagasy survey, we identify the effect of maternal cognitive skills by gradually introducing control variables that can proxy the unobserved variables and child's heterogeneity. Ultimately, we estimate the reduced demand child health as follows

$$H_{ij} = a + \partial CHij + \beta' MCog_{ij} + \rho' MX_{ij} + \varphi' Aij + \pi' Cij + NonCog_{ij} + u_{ij} \quad (1)$$

Where *Hij* is the height-for-age score of child *i* from mother *j*. $\boldsymbol{MCog_{ij}}$ is the mothers' standardized total score, sum of the three test scores Math, French, Life Skills.[76]We also estimate our models with a factor of the three scores and the results are maintained.[77] $\boldsymbol{CH_{ij}}$ is a vector of child's age cohort dummies and gender.

[75] Behrman et al (2009), using the 35 year- panel data collected to the evaluate a nutritional interventional intervention in Guatemala, also use mother's parents education to instrument maternal schooling but in this case they also have other community-level instruments such as a dummy variable if the mother was exposed to this nutritional intervention when she was 36 months.

[76] We estimate the models with each of these test scores and the sum of Math and French test scores but they are not statistically significant.

[77] Results are upon request.

MX_{ij} is a vector of mother's socioeconomic characteristics including: i) height which has been a widely used as an indicator of child health endowments, ii) parents' education (i.e., child grand-parents), iii) the 2004 number of siblings and their average education, when the young mothers were in average 13 to 16 years old, and iv) grandparents' education (i.e., child great-grandparents), information that comes from the 2004 round of the survey.[78] These intergenerational variables are related to mother's education and cognitive ability and allow us to control for the young women's human capital endowments (e.g., genetics, ability). To complete this set of childhood socioeconomic characteristics, from the 2004 school survey, we include two school facility indexes related to the primary school infrastructure and teachers' quality.[79] We estimate the models with and without years of schooling to evaluate their relative importance in children's health. It is possible that years of schooling might be endogenous to the tests scores. Nevertheless, studies in Madagascar and Senegal on the determinants of cognitive skills have not found enough evidence to reject the null hypothesis of exogeneity (Glick et al., 2009).

To disentangle the effect of income from cognitive skills and given that the Malagasy data does not include information related to household income and

---

[78] We also estimate the models with alternate specifications using the number of sisters and their average education but we do not find any different results.

[79] Using factor analysis, the school facility index uses variables related to school infrastructure such as availability of bathrooms, blackboards, and electricity among other variables, while the school teaching index uses number of experience and education of teachers and principals.

expenditures, the $\boldsymbol{Aij}$ vector includes a 2004 asset index and young women's spouse education as a proxy of a household wealth.[80]

We also rely on a set of community variables that help address the heterogeneity in children's health production function at the community level. $\boldsymbol{Cij}$ is a vector of community variables related to social infrastructure such as access to health centers, piped water, and electricity. Additionally, from the 2001 community census, we include a remoteness index which is created using factor analysis and information on community distances to the main social infrastructure services and transportation.[81] In combination, these control variables allow us to uncover some of the unobservable characteristics at the community level that might be related to a child's health. As a robustness check, we also estimate our models with community fixed effects.

In addition, we analyze if the gradient of maternal cognitive skills and child health is affected by the mothers' non-cognitive skills. Non-cognitive ability refers to a broad spectrum of socio-emotional status, personality traits, time preferences, among other features. In this paper, we use a measure of maternal non-cognitive skills, her Big Five personality traits, represented by vector $\boldsymbol{NonCog_{ij}}$ in equation 1. Using a special module of the Malagasy panel data and factor analysis, these personality traits are constructed by Villa and Sahn (2014) and defined as follows:[82] i) *Openness to Experience* is the degree to which a person is curious, needs intellectual stimulation,

[80] See Glick and Sahn (2004) for details of the asset index construction

[81] The remoteness index includes health services, banks, post offices, schools, taxis, courts, markets, inputs, extension services, and veterinarians as well as access to national and provincial roads, utilities, media and other markets and several measures of access to transport.

[82] Villa and Sahn model the Big Five Personality Traits as unobserved variables and use factor analysis to uncover their latent distributions. Please see Table 1 in Villa and Sahn (2014) for the variables used in the survey to construct these personality traits.

change and variety; ii) *Conscientiousness* captures the trait of being hardworking, organized and dependable as opposed to lazy, disorganized and unreliable; iii) *Extroversion* captures the preference for human contact, empathy, gregariousness, assertiveness and a wish to inspire people; iv) *Agreeableness* is the degree to which someone is cooperative, altruistic, modest, warm and agreeable vs. cold, disagreeable and antagonistic and v) *Neuroticism* is the extent to which an individual is insecure, anxious, depressed and emotional rather than calm and self-confident. By adding these personality traits, we expect to capture unobserved variables that are related to the mother's cognitive ability. These Big Five personality traits are assumed to be innate, and thus, they are not endogenous to other mothers' socioeconomic characteristics. We also estimate the models including the interaction between each of the mother's personality traits and her standardized total score. Table 3.1 summarizes all the child, mother, and community characteristics used in the estimation.

3.4 Results and Discussion

Table 3.2 presents the models of child's health where we focus on the impact of maternal cognitive skills. We begin with the most parsimonious model and gradually add control variables that might be correlated with the mother's cognitive skills and her child's health.

Using the results shown in the first column, we find that one standard deviation increase in the mother's cognitive test score increases her child's height-for-age by 0.28 standard deviations; this coefficient is statistically significant at the 5% level. Adding to this basic model other maternal human capital variables such as the mother's height and years of schooling has little impact on the test score coefficient. In the second column of Table 3.2 we actually leave out the test score variable and in that case the mother's years of schooling does have a positive effect (p-value 0.11), but as noted above, it vanishes when including mother's cognitive skills, much of which is acquired at school. Like cognition, the taller the mothers the less stunted their children, a result that has been well established in the empirical evidence for developing countries (Thomas et al, 1991; Rubacalva, 2004; Federov and Sahn, 2005; Behrman et al, 2009). In fact, according to the results of column 3, 1onepercent increase in the mother's height (or 1.5 cm over the sample mean) increases her child's height-for-age by 0.05 standard deviations.

As we emphasize above, mother's cognitive skills might well capture part of her childhood human capital endowment such as parental transmission of knowledge, practice and values. In addition, cognitive skills are presumably not just a function of the amount of schooling, but its quality as well. Using information gathered on the young mothers when they were between 13 to 16 years old, we include a range of additional controls for earlier life experiences and circumstances in column 4 of Table 3.2: i) parents' education (child's grandparents), ii) number of siblings and their average education, iii) grand parents' education (child's great grandparents), and iv) two 2004 indexes of school quality of the closest primary school in the communities were the mothers' were residing. We observe that the point estimate of the total standardized score increases to 0.34 and remains statistically significant at the 5% level. This result suggests that the effect of mother's cognitive ability on child health is independent of her parents' education, which might capture parental transmission of knowledge and values. This effect is also independent of the childhood characteristics of the location where she grew up, including the quality of the local primary school. Neglecting these childhood circumstances would thus have negatively biased the returns to cognitive ability return on child health. This finding is consistent with the results of Rubacalva et al. (2004) in Mexico. In this case, the mother's cognitive ability, measured by the Raven test, remains statistically significant when mother's childhood background such as parents' education and childhood community characteristics are included as covariates.

Given the possibility that cognitive test scores can also have an income effect in children's health production, we also include a 2004 asset index and the mother's spouse education in the regressions in the next set of models to see if income is the pathway through which the mother's human capital is affecting child nutrition. In column 6 of Table 3.2, we observe that the point estimate of the total standardized score slightly decreases, but remains statistically significant. Similar results are found when the models are estimated with the 2012 asset index. However, we prefer the specification with the lagged asset index since the 2012 one might have a problem of reverse causality.

Table 3.2. Effect of Mother's Cognitive Skills on Children Height-for-age (HAZ)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ	HAZ –CFE
Mother Total Standardized Score	**0.283****		**0.260***	**0.341****	**0.301***	**0.314****	**0.333****	**0.318***	**0.474****
	[0.110]		**[0.147]**	**[0.149]**	**[0.163]**	**[0.155]**	**[0.158]**	**[0.170]**	**[0.193]**
Missing Total Standardized Score	0.870		0.961	0.983	1.006	0.958	1.022	1.029	2.249**
	[0.753]		[0.752]	[0.769]	[0.768]	[0.763]	[0.770]	[0.771]	[1.110]
Mother's Years of Schooling		0.0531	-0.00990		0.0287			0.0124	-0.00999
		[0.0338]	[0.0448]		[0.0495]			[0.0498]	[0.0604]
Mothers ln height			5.625*	5.903**	5.829**	5.896**	5.100*	5.085*	5.222*
			[2.874]	[2.943]	[2.947]	[2.984]	[2.998]	[3.002]	[2.986]
Grandmothers Education				-0.0291	-0.0331	-0.0295	-0.0364	-0.0378	-0.0879*
				[0.0418]	[0.0432]	[0.0424]	[0.0442]	[0.0454]	[0.0503]
Grandparents Education				-0.0176	-0.0226	-0.0155	-0.0161	-0.0177	-0.00883
				[0.0370]	[0.0375]	[0.0385]	[0.0389]	[0.0391]	[0.0399]
Great Parents Education (Max)				-0.0181	-0.0180	-0.0212	-0.0525	-0.0517	0.00934
				[0.122]	[0.122]	[0.124]	[0.122]	[0.122]	[0.145]
Missing Great parents education				0.175	0.162	0.189	0.0656	0.0619	-0.0462
				[0.350]	[0.353]	[0.359]	[0.366]	[0.367]	[0.400]
No of siblings				-0.0104	-0.00739	-0.00315	0.00222	0.00288	0.0271
				[0.0550]	[0.0551]	[0.0563]	[0.0561]	[0.0561]	[0.0607]
Siblings average education				-0.0742	-0.0815	-0.0754	-0.0632	-0.0656	-0.0810
				[0.0548]	[0.0560]	[0.0551]	[0.0564]	[0.0576]	[0.0712]
2004 School Facility Index				0.0544	0.0587	0.0712	0.0279	0.0303	
				[0.178]	[0.178]	[0.184]	[0.229]	[0.229]	

Table 3.2 (Continued)

2004 School Teaching Index				0.230	0.228	0.214	0.139	0.139	
				[0.242]	[0.242]	[0.245]	[0.261]	[0.262]	
2004 Asset Index						-0.0792	-0.0986	-0.103	-0.104
						[0.160]	[0.191]	[0.192]	[0.257]
Spouse Education						0.0279	0.0203	0.0175	-0.00475
						[0.0398]	[0.0393]	[0.0397]	[0.0462]
Missing Spouse Education						-0.0171	-0.0670	-0.0687	-0.319
						[0.251]	[0.255]	[0.255]	[0.279]
Access Community Health Center							0.729**	0.724**	
							[0.359]	[0.358]	
Urban (Y=1)							1.012*	1.007*	
							[0.587]	[0.585]	
Access to Piped Water (Y=1)							-0.117	-0.118	
							[0.260]	[0.261]	
Access to Electricity (Y=1)							-0.0856	-0.0812	
							[0.321]	[0.323]	
2001 Remoteness Index							0.388	0.371	
							[0.460]	[0.467]	
2001 Remoteness Index quadratic							-0.0669	-0.0640	
							[0.0773]	[0.0786]	
Child gender and age cohort dummies	X	X	X	X	X	X	X	X	X
Mother's Age Cohort Dummies	X	X	X	X	X	X	X	X	X
N	500	500	495	482	482	482	482	482	489
R-sq	0.037	0.022	0.045	0.053	0.053	0.054	0.065	0.066	0.074

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses.

We also include community variables relevant to the child health's production function such as access to community health center, piped water, electricity and whether the community is urban or not. Also, to address the potential omitted variable bias at the community level, we add a 2001 remoteness index which is a good indicator of early economic and social development at the community level. Column 7 of Table 3.2 shows that having access to community health center or living in an urban area increases the child's height-for-age by 0.7 and one standard deviation, respectively. Consistent with empirical evidence, this result underlines the importance of health infrastructure in the long-term child health outcomes in Madagascar (Umapathi, 2008). The point estimate of the mothers' cognitive skills is once again robust to the inclusion of these community variables. As a robustness check given the possibility that other community variables related to health input and prices may be omitted, we estimate our model using community fixed effects in column 9 of Table 3.2. The return of maternal cognitive ability on child health is robust to this specification: the point estimate of the total score increases to 0.47 but it is still in the 95% confidence interval compared to the estimate in columns 7 and 8. Different from Thomas et al. (1991) and Alderman et al. (2004), we do not find any statistical effect of the interaction of the mother's total score and the community health services.[83] Thus, we cannot infer whether the cognitive skills increase mothers' use of the health services available in the community.

The results in our preferred specifications (columns 7 and 9 in Table 3.2) indicate that one standard deviation of the mother's standardized total score can

[83] These results are upon request

increase her child's height-for-age between 0.33 and 0.47 standard deviation. This magnitude is relevant in terms of public policy when it is compared to the interventions that aim to decrease malnutrition in Madagascar. For instance, Umapati (2008) reports that Seecaline, the largest Malagasy nutritional program, increased the child height-for-age by 0.32 standard deviation during the period 1999 and 2007, only for those children under age of 5 whose mothers have at least secondary school. In our sample, only 7% of the mothers have completed secondary school. Indeed, Herrera and Sahn (2013) show that these young mothers have lower school attainment and cognitive skills than their non-mother counterparts in the cohort. Therefore, our results suggest that even at lower levels of schooling (i.e., primary school) maternal cognitive skills can have a positive return on child health.

Our findings therefore underline the importance of cognitive skills on the health of children, and that the impact of educational attainment is operating through the impact on cognition. This evidence is consistent with Behrman et al. (2009) who find that maternal cognitive skills, measured by a reading and Raven test scores, tend to be more predictive than school attainment in explaining children's human biological capital (birth weight and anthropometrics measures) in Guatemala. Nevertheless, this finding differs from other studies in developing countries. Glewwe (2009) , using data from Morocco, argues that there is no direct effect of mothers' numeracy and literacy skills on child health, but there is a positive effect through the health knowledge that is not directly learned in school but rather acquired through literacy and numeracy skills.

Unfortunately, we lack information to identify if these cognitive abilities are translated into better health knowledge or change in related health behaviors such as reproductive or child care practices. We test the hypothesis of a differentiated effect of the mother's total score by birth order, by exploring if there is a gain in child rearing experience after the first child. We do not find statistically significant effects of the interaction of the mothers' total score and age cohort dummies or birth-order variables.

3.4.1 The role of Non-Cognitive Skills

Empirical evidence, mostly from the US, has shown that non-cognitive skills have a role as well as relative importance compared to cognitive skills in determining economic outcomes such as school attainment, risky and health behaviors and labor market participation, among others (Heckman et al. 2006). Empirical evidence on the effect of personality traits on health outcomes and behaviors is scarce, particularly in the context of developing countries. Empirical evidence from the field of psychology and health sciences has shown that personality traits such as conscientiousness, openness to experience, and agreeableness generally have a positive effect on health outcomes such as longevity (Almlund et al., 2011). In the economics research, Heckman et al. (2006) show in the US that non-cognitive abilities such as locus of control and self-esteem, affect risky behaviors among youth such as the probability of smoking, using marijuana, and being a teenager mother. Similar results are found by Mendiola and Walker (2013) in Britain; individuals with external locus control, low self-esteem or low levels of conscientiousness are more likely to engage in risky health behaviors. This evidence is not conclusive. For instance, Cutler and Lleras-

Muney (2010) show that individual's risk aversion and sense of control do not explain the gradient between education and health behaviors in the US and Britain.

Most of these studies analyze the effect of an individual's personality traits on his/ her own later youth and adulthood health behaviors. To our knowledge there is no empirical evidence that point out the role of maternal personality traits on her child's height-for-age. We next examine if the impact of a mother's cognitive ability on her child's health is affected by adding information on her Big Five personality traits, into the model.

Table 3.3 shows first the initial model of child health on maternal cognitive skills including all the covariates and gradually includes the Big Five personality traits as well as their interaction with the test score. The results show that these non-cognitive skills do not have a statistically significant effect on children's height-for-age. The point estimate of the mother's test score remains statistically significant and does not change in magnitude. In, column 2 we interact these personality traits with the test score and find that the interaction with *Openness to Experience* is positive and statistically significant. This finding suggests that mother's return of cognitive skills is reinforced by certain aspects of her personality such as curiosity and need for intellectual stimulation. The fact that none of the five personality traits have a statistically significant effect on child's height-for-age is surprising in light of the theoretical framework of Cuhna and Heckman (2010) where maternal personality traits can have a role in determining children's capabilities, including health. Nevertheless, these results should be interpreted only in the context of our sample of Malagasy young mothers. It might be possible that there is not enough variance of these traits among the mothers' personality traits to identify statistically significant effects. In addition, the analysis of this paper is limited to the Big Five personality traits, but it does not imply that other measures of maternal non-cognitive traits (i.e., depression) might have a statistically significant association with child health. Further research should explore if other measures of the mother's non cognitive skills have a role in child height-for-age and other health outcomes such as birth weight in the context of developing countries.

Table 3.3 Mother's Cognitive and Non-Cognitive Skills Return on Child Height-for-Age

	(1)	(2)	(3)	(4)
	HAZ	HAZ	HAZ	HAZ
Mother Standardized Total Score	**0.361****	**0.332***	**0.474*****	**0.445*****
	[0.162]	**[0.170]**	**[0.160]**	**[0.171]**
Missing Total Standardized Score	1.226	1.242	1.262	1.276
	[0.805]	[0.806]	[0.798]	[0.798]
Mothers' Years of Schooling		0.0239		0.0219
		[0.0501]		[0.0499]
Openness	0.100	0.103	0.100	0.103
	[0.174]	[0.173]	[0.174]	[0.174]
Agreeableness	0.0725	0.0766	0.0733	0.0773
	[0.216]	[0.216]	[0.205]	[0.205]
Conscientiousness	-0.0729	-0.0821	-0.0943	-0.103
	[0.214]	[0.213]	[0.211]	[0.210]
Extroversion	-0.302	-0.307	-0.261	-0.267
	[0.236]	[0.236]	[0.235]	[0.235]
Neuroticism	-0.0530	-0.0564	-0.00515	-0.00814
	[0.118]	[0.117]	[0.116]	[0.116]
ZscoreTotal*Openness			0.451**	0.447**
			[0.183]	[0.184]
ZScoreTotal *Agreeableness			0.343	0.343
			[0.228]	[0.229]
ZScoreTotal*Conscientiousness			-0.0821	-0.0808
			[0.194]	[0.194]
ZScoreTotal*Extroversion			-0.291	-0.291
			[0.259]	[0.259]
ZScoreTotal*Neuroticism			0.187	0.190
			[0.120]	[0.120]
Child gender and age cohort dummies	X	X	X	X
Mothers' Socioeconomic Characteristics	X	X	X	X
Community Variables	X	X	X	X
N	481	481	481	481
R-sq	0.072	0.073	0.104	0.105

Notes: ***, **, *: significant at 1%, 5%, and 10% levels respectively. Robust Standard errors reported in parentheses. Mother's socioeconomic characteristics include: mother's height, 2004 asset index, spouse's education, parents and great grandparents education, no. of siblings and their average education.

3.5 Conclusion

Empirical studies have shown the importance of mother's human capital on determining children health outcomes in developing countries. Nevertheless, the role of maternal cognitive ability in child long term health outcomes has been less explored, particularly, in the Sub-Saharan African context. We contribute to this gap in the literature by addressing whether the mother's cognitive skills, measured by the standardized sum of French, Math and Life skills test scores, affect her child's height-for-age in Madagascar, a country with the sixth highest rate of malnutrition in the world.

Using a panel data survey that follows a sample of young mothers from their childhood, we address the potential endogeneity between maternal education and children health by gradually adding a large set of mother's socioeconomic and intergenerational variables as well as community characteristics to the OLS models of maternal cognitive skills on child Z-score height-for-age. We control for mothers' years of schooling and height, and an extensive set of mother's childhood socioeconomic background such as her parents', grandparents' and siblings' education as well as the primary school quality indexes of the community where the mother lived during childhood. One unique feature of the data allows us to control for the mother's non-cognitive skills, measured by her Big Five personality traits.

Our findings point out that the relationship between mother's years of schooling and child health is mediated by her cognitive skills. One standard deviation in the mother's standardized total score increases her child's height-for-age by 0.31 and 0.44 standard deviations in our preferred specifications. Our results also show that the effect of maternal cognitive ability is not channeled through an income effect or better use of health infrastructure. We lack information to identify if these maternal cognitive abilities are translated into better health knowledge or change in related health behaviors such as reproductive or child care practices. These findings are robust to the inclusion of the mother's Big Five personality traits. Further research should analyze other measures of maternal non-cognitive skills to establish if the gradient of maternal education and child health is independent of the mother's non-cognitive ability, specially, in the context of developing countries.

These results suggest that policies that aim to increase young women's human capital, particularly, their cognitive skills such as numeracy and literacy abilities, can be translated into better health outcomes of their children breaking potential channels of intergenerational transmission of poverty.

REFERENCES

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. D. (2011). Personality psychology and economics (No. w16822). National Bureau of Economic Research.

Aslam, M., & Kingdon, G. G. (2012). Parental Education and Child Health Understanding the Pathways of Impact in Pakistan. *World Development*, 40(10), 2014-2032.

Behrman, J. R., Murphy, A., Quisumbing, A. R., & Yount, K. (2009). Are Returns to Mothers' Human Capital Realized in the Next Generation?. Free downloads from IFPRI.

Block, S. A. (2007). Maternal nutrition knowledge versus schooling as determinants of child micronutrient status. *Oxford Economic Papers*, *59*(2), 330-353.

Christiaensen, L., and Alderman, H. (2004). Child Malnutrition in Ethiopia: Can Maternal Knowledge Augment the Role of Income?. *Economic Development and Cultural Change*, *52*(2), 287-312.

Cuhna, F. and Heckman, J (2007). The Technology of Skill Formation. *American Economic Review* 97 (2):31-47.

Cutler, D. and A. Lleras-Muney (2010): Understanding differences in health behaviors by education. *Journal of Health Economics*, 29, 128.

Galasso, E., & Umapathi, N. (2009). Improving nutritional status through behavioral change: lessons from Madagascar. *Journal of Development Effectiveness*, *1*(1), 60-85.

Güneş, P. M. (2013). The Role of Maternal Education in Child Health: Evidence from a Compulsory Schooling Law.

Fedorov, L. and Sahn, D. (2005). Socioeconomic determinants of children's health in Russia: A longitudinal study. *Economic Development and Cultural Change* 53.2 (2005): 479-500.

Glewwe, P. (1999) Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco. *The Journal of Human Resources*. Vol. 34, No. 1 pp. 124-159

Glick, P., Randrianarisoa, J. C., & Sahn, D. E. (2011). Family background, school characteristics, and children's cognitive achievement in Madagascar. *Education Economics*, *19*(4), 363-396.

Heckman, J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children, *Science*, 312: 1900-1902.

Herrera, C. and Sahn, D (2013).The Impact of Early Childbearing on Schooling and Cognitive Skills among Young Women in Madagascar. *Cornell University Food and Nutrition Policy Program Working Paper* 247 (2013).

Hoddinott, J., Maluccio, J. A., Behrman, J. R., Flores, R., & Martorell, R. (2008). Effect of a nutrition intervention during early childhood on economic productivity in Guatemalan adults. *The Lancet*, *371*(9610), 411-416.

Kovsted, J., Pörtner, C. C., & Tarp, F. (2002). Child health and mortality: does health knowledge matter?. *Journal of African Economies*, *11*(4), 542-560.

McCrary, J., & Royer, H. (2006). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American Economic Review*, 101(1): 158–9

Mendolia, S. and Walker, I.(2013).The Effect of Non-Cognitive Traits on Health Behaviours in Adolescence. IZA Discussion Paper No. 7301. Available at SSRN: http://ssrn.com/abstract=2245971

Rubalcava, L.N., Teruel, G.M., (2004). The role of maternal cognitive ability on child health. *Economics and Human Biology* 2 (3), 439–455

Sharp, M., and Kruse, I. (2011). Health, Nutrition, and Population in Madagascar, 2000-09 (No. 216). World Bank Publications-free PDF.

Strauss, J.and Thomas D. (1995) "Human Resources: Empirical Modeling of Household and Family Decisions." in Jere Behrman and T.N. Srinivasan, eds., *Handbook of Development Economics*, vol. 9, 1883-2023. New York: Elsevier Science.

Umapathi, N. (2008). Maternal education, child-care and nutritional status: Lessons from a nutritional program.

Villa, K. and Sahn, D.(2014).The Role of Personality and Cognition in Determining Occupational Choice and Earnings. Dissertation, Cornell University.

World Health Organization (1983). *Measuring Change in Nutritional Status: Guidelines for Assessing the Nutritional Impact of Supplementary Feeding Programs for Vulnerable Groups*. Geneva: World Health Organization.